ESD-TR-66-449

Technical Note

1966-50

Signal Processing Characteristics of the Peripheral Auditory System

T. J. Goblick, Jr.
R. R. Pfeiffer

30 September 1966

Prepared under Electronic Systems Division Contract AF 19(628)-5167 by

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Lexington, Massachusetts

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the U.S. Air Force under Contract AF 19(628)-5167.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

SIGNAL PROCESSING CHARACTERISTICS OF THE PERIPHERAL AUDITORY SYSTEM

T. J. GOBLICK, JR.

R. R. PFEIFFER

Group 66

TECHNICAL NOTE 1966-50

30 SEPTEMBER 1966

LEXINGTON MASSACHUSETTS

ABSTRACT

The fundamental question in speech compression is that of determining the minimum information rate that must be maintained between speaker and listener in order to achieve a specified level of speech fidelity or quality. The problem in answering this question is that a measure of speech quality must first be defined. Any meaningful definition of speech quality clearly must consider the manner in which speech is processed by the listener. If the details of the signal processing in the auditory system were known, speech quality could be defined in terms of the sensitivity of the listener to distortions of signals within the auditory system.

A study of the manner in which sounds are processed by the human auditory system was done to provide the basic information to define a measure of speech fidelity. The mechanical or sound-conducting parts of the auditory system are reasonably well understood and can be considered as linear systems in an engineering sense. The neural processing of the peripheral auditory system is only partly understood. Further experimental work augmented with computer simulation of neural models is necessary to fully understand the nature of the neural data processing.

Accepted for the Air Force
Franklin C. Hudson
Chief, Lincoln Laboratory Office

Table of Contents

SIGNAL PROCESSING CHARACTERISTICS OF THE PERIPHERAL AUDITORY SYSTEM

Thomas J. Goblick, Jr.
Russell R. Pfeiffer*

CHAPTER I

Introduction

Speech is a very convenient and rather efficient means of person-to-person communication. Except for very short distances, a special communication channel is needed to transmit speech from speaker to listener, thus putting the study of the speech-generation process within the domain of the communication engineer. For various reasons, speech is often represented in digital form for transmission over a channel. One reason is that digital speech is easily coded for privacy. Another is that the speech coding and digital channel coding equipments are interfaced very conveniently, allowing the speech coding equipment to be designed independently of the channel.

The cost of channel capacity has motivated a large effort to minimize the digital data rate that must be sent over a channel to achieve satisfactory speech communication. Obviously the minimum channel capacity required by a digital speech communication system depends on the desired speech quality. Therefore, the fundamental question is that of finding the minimum channel capacity consistent with a given level of speech quality. The answer to this question, of course, lies in determining how the various features of speech

* Presently at the Department of Electrical Engineering, Washington University, St. Louis, Missouri.

contribute to speech quality, which in turn is totally dependent upon the data processing performed by the listener. Therefore, the human auditory system is also, or should be, of interest to the communication engineer.

For example, let us take the electrical output of a microphone as our representation of speech and suppose we ignore all knowledge of the auditory system of the listener. Then we must digitize speech purely on the basis of preserving the fidelity of the microphone output (from which speech sounds can later be recovered by a loudspeaker). This has been done and it is now well known that this approach requires excessively high digital data rates. But suppose two sounds give rise to identical firing patterns in each of the neurons in the auditory nerves of a listener. Regardless of the form of the subsequent processing of this neural representation of speech at higher levels of the nervous system and brain, the two sounds must be perceived as being identical. * In fact, the same sound heard twice in succession will not give rise to identical neural firing patterns due to the internal noise of the auditory system. Therefore, the neural representations of different sounds need not correspond to exactly the same neural firing patterns in order to be perceived as the same sound.

It would be meaningful to digitize speech to provide some level of fidelity of the neural firing patterns in the auditory nerve, if the relation between the fidelity of neural firing patterns and speech quality is known. This viewpoint could be clearly generalized to preserve the form that speech assumes at

---

* All other things being equal, of course; that is, the distractions due to other stimuli than acoustic, and the mental state of the individual should be the same.

some other level within the auditory system than the auditory nerve. Some of the various forms of speech within the peripheral auditory system are:

(a) eardrum vibrations;

(b) mechanical motions of the middle ear bones; and

(c) vibratory patterns of the basilar membrane within the cochlea.

As speech signals are traced further into the auditory system, more of the characteristics and constraints of the system are taken into account. Achieving satisfactory speech quality by preserving the fidelity of the signals at some level within the auditory system cannot require a higher data rate than that required for the same speech quality if one considers signals at a more peripheral level, since satisfactory signals at the peripheral level must result in signals that are satisfactory at all following levels. Therefore, the deeper one probes into the auditory system, the lower the resulting data rate ought to be to provide satisfactory speech quality. It is our lack of knowledge of the human auditory system that limits the extent to which this viewpoint can be carried.

We take the view that a precise description of speech quality must be derived from the auditory system of the listener. This report summarizes an effort to determine quite precisely the extent to which the processing of auditory stimuli in the intensity and frequency range of speech sounds by the human auditory system is understood. * This study was conducted with the goals of modeling and simulating the various parts of the human auditory system in mind, since,if parts of the system could be simulated with results that agreed

---

* Such topics as auditory thresholds, binaural localization, etc., are, therefore, not included in this study.

with all known experimental data, that aspect of the system is presumably well understood. Modeling is necessary to compactly characterize the auditory system processing for mathematical analysis as well as computer simulation anyway. Simulation is also desirable in order to study experimentally the sensitivity of the auditory system to the various features of speech. It is our hope that this effort will lead to an objective measure of speech fidelity which will be of value in the analysis of mathematical models of speech communication systems as well as in the evaluation of performance of existing systems.

In Chapter II of this report, we present a brief review of the anatomy of the peripheral auditory system of humans. Thereafter, the linear, or sound-conducting (Chapter III), and non-linear, or neural (Chapter IV), stages of the auditory system are discussed along with some of the past efforts at modeling various auditory system functions at these levels. Some of our own efforts at modeling are also included. Chapter V discusses the few attempts to obtain an overall model of the auditory system and the shortcomings of these models. This chapter also concludes with a summary of the state of our knowledge of the peripheral auditory system of humans with respect to speech inputs, and discusses some of the available means of extending our knowledge.

Communications Theory References

R. M. Fano, "The Information Theory Point of View in Speech Communication," J. Acoust. Soc. Am. 22 (November 1950).

R. M. Fano, Transmission of Information (M. I. T. Press-Wiley, 1961).

J. L. Flanagan, Speech Analysis, Synthesis, and Perception (Academic Press, 1965).

T. J. Goblick, Jr., J. L. Holsinger, "Analog Source Digitization: A Comparison of Theory and Practice," to be published in IEEE Trans. on Information Theory.

C. E. Shannon, "A Mathematical Theory of Communication," Bell Syst. Tech. J. 27, 379, 623 (1948).

C. E. Shannon, "Coding Theorems for a Discrete Source with a Fidelity Criterion," IRE Nat'l Convention Record, Part 4, 142 (1959).

J. M. Wozencraft, I. M. Jacobs, Principles of Communication Engineering (Wiley, 1965).

CHAPTER II

Some Details of Structure and Anatomy of the Auditory System

This chapter deals primarily with the description of the anatomical structure of certain parts of the auditory system. We will restrict ourselves to those aspects of the system that we feel are directly pertinent to the purposes of our study, i.e., the processing of speech. We have separated the system into simpler subsystems which can first be studied by themselves. Figure 2-1 is a block diagram showing our subdivision of the system. A discussion of each subsystem follows, starting with the most peripheral.

The external ear consists of the pinna (auricle) and auditory canal (external auditory meatus); the middle ear consists of the ossicles and a system of hollow cavities; the inner ear consists of the cochlea and the cochlear (auditory) nerve; and finally the central auditory nervous system consists of those portions of the system located in the brain and brain stem.

The peripheral portion of the system, for our purposes, consists of the external, middle, and inner ears. These structures will receive most of our attention. Fortunately, those peripheral parts of the system are among those which are, at present, best understood. This latter fact is due to

(a) the relative simplicity of these parts as compared to the higher structures;

(b) the availability of detailed, systematic, experimental results; and

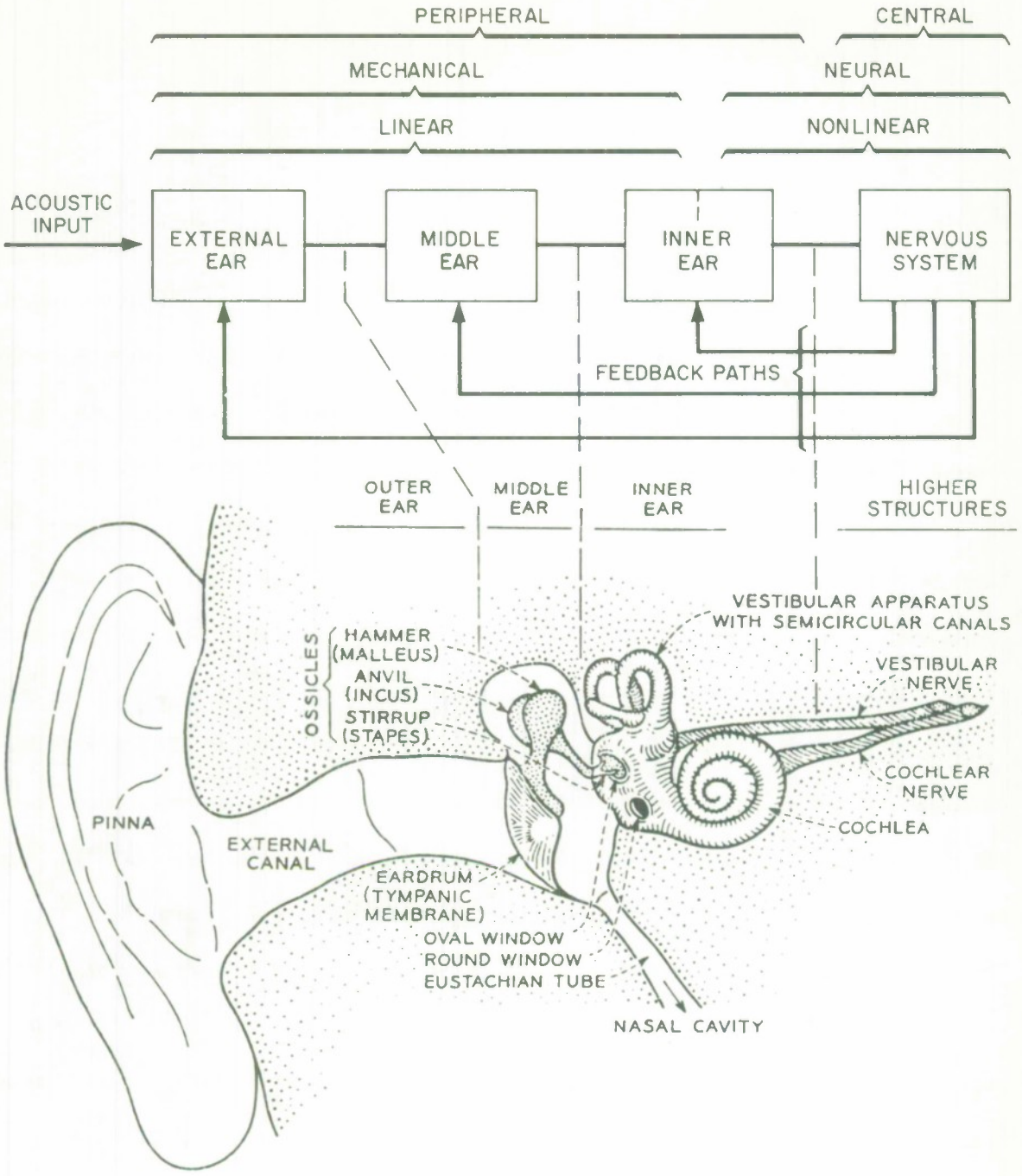

Fig. 2-1. Block diagram of peripheral auditory system. From Weiner (1949).

(c) the linearity (in an engineering sense) of some of these parts over the dynamic ranges of interest to us.

The function of these peripheral blocks will be discussed later.

The External Ear

The external ear is strictly an acoustic device. The pinna, or auricle, is a complicated cartilaginous structure, but despite its complexity it does not appear to serve a major function in hearing in humans (probably due to its small size relative to that of the head). The auditory canal, on the other hand, does have some effect on the signals that reach the eardrum (see below). This canal also is a cartilaginous structure--an approximately cylindrical tube of about 2.3 cm in length.

The feedback paths to the external ear (Fig. 2-1) represent the control of the muscle actions which orient the pinna. These actions can be safely ignored in man. In other species, however, the pinnae are important, the auditory canals are more complicated than in man, and the reflex action cannot be completely ignored.

The Middle Ear

The middle ear is a system of small bones (the well-known hammer, anvil, and stirrup) and their supporting ligaments and muscles. These bones serve to transmit acoustic vibrations from the tympanic membrane (eardrum) to the cochlea. Figure 2-2 is a view of the middle ear of man from the inside--a view from the right of the sketch in Fig. 2-1. It can be seen that the malleus is attached to the eardrum and is further supported by (1) the

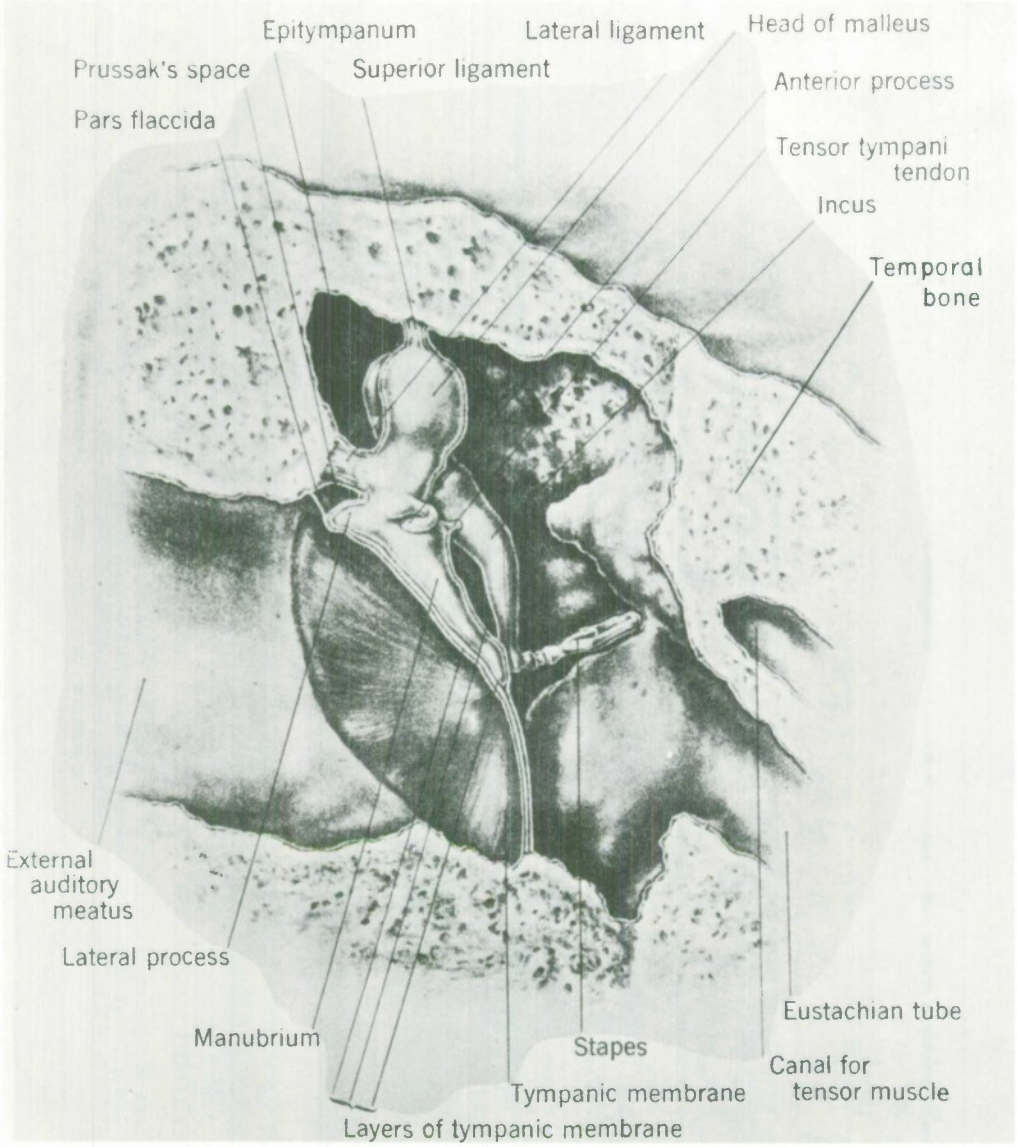

Fig. 2-2. The middle ear on the right side, seen from the front. The tensor tympani muscle and most of its tendon have been removed. From John B. Deaver, Surgical Anatomy of the Human Body (The Blakiston Company, 1926).

superior ligament, (2) the tensor tympanic muscle, and (3) its attachment to the incus. The incus, in addition to being attached to the malleus, is supported by its attachment to the stapes and the posterior ligament. The stapes, in addition to being attached to the incus, is further supported by the stapedius muscle. The base of the stapes (the footplate), which is viewed directly in Fig. 2-3, covers the oval window of the cochlea. The feedback paths shown in Fig. 2-1 represent the central controls of the tensor tympani and stapedius muscles.

The Inner Ear

The major component of the inner ear is the cochlea and its contents. The cochlea is a bony, helical structure consisting of a tubular canal spiraled around a central core, the modiolus (Fig. 2-3). Within this helical canal, and running the length of it, is a very intricate structure named the organ of Corti. Some details of the inside of the canal are shown in a sketch of its cross section (Fig. 2-4). The organ of Corti is supported between the spiral ligament on the outer wall of the canal and the bony septum on the inner wall. That portion of the canal below the organ of Corti is called the scala tympani; that portion above the organ of Corti is divided into two scalae called the scala media and scala vestibuli. Each of these scalae is filled with fluid. The scala tympani and the scala vestibuli are joined together at the very apex of the cochlea; whereas, at the base of the cochlea they are terminated by the round and oval windows respectively (see Fig. 2-1). The fluid system of the scala media is not connected with that of the other scalae.

The intricacies of the organ of Corti are amplified in Fig. 2-5. One of the main features of the organ which is of interest to us are the hair cells.

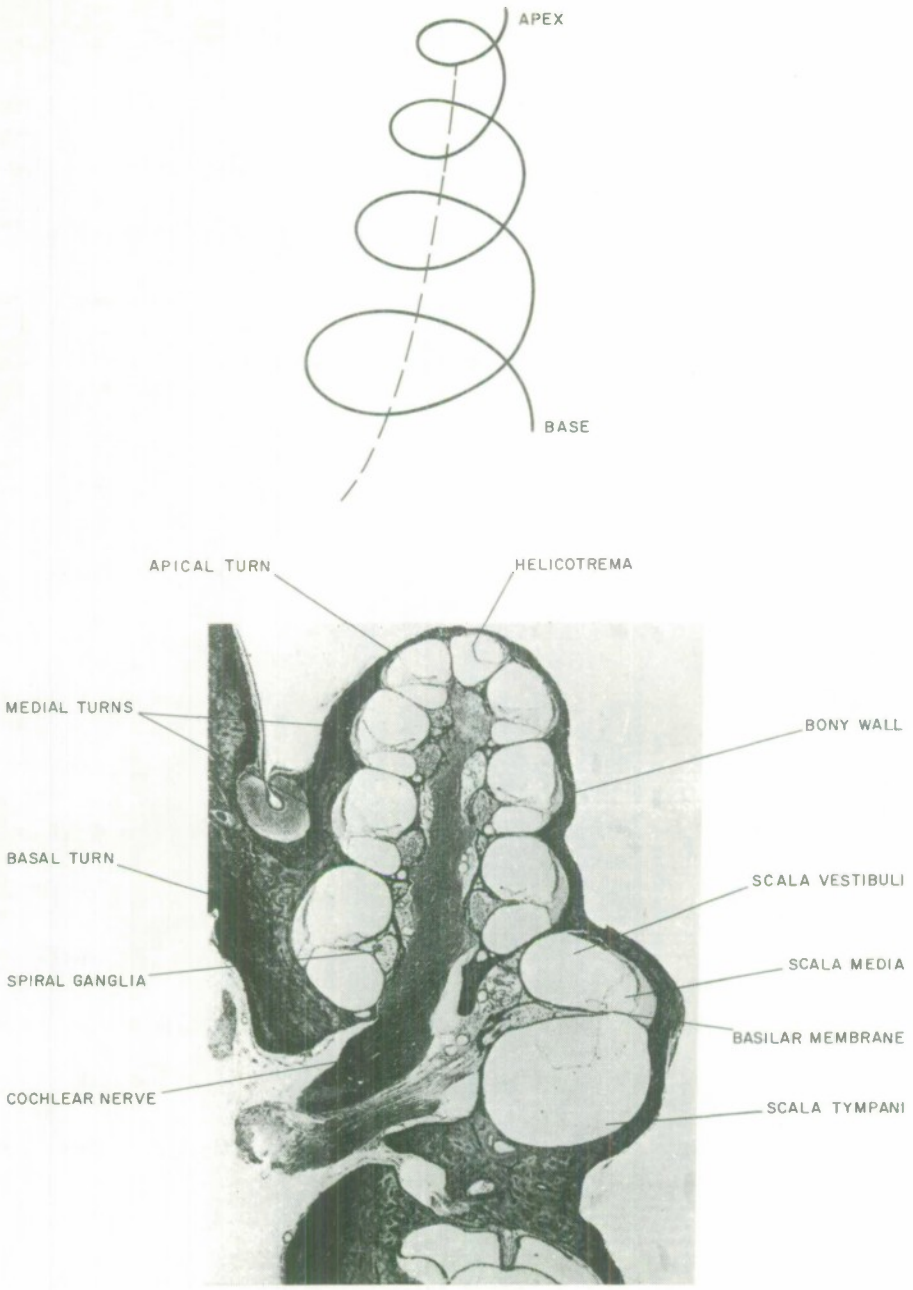

Fig. 2-3. The cochlea of a guinea pig is similar to that of a human except it has more turns and is somewhat more helical. From Davis (1961).

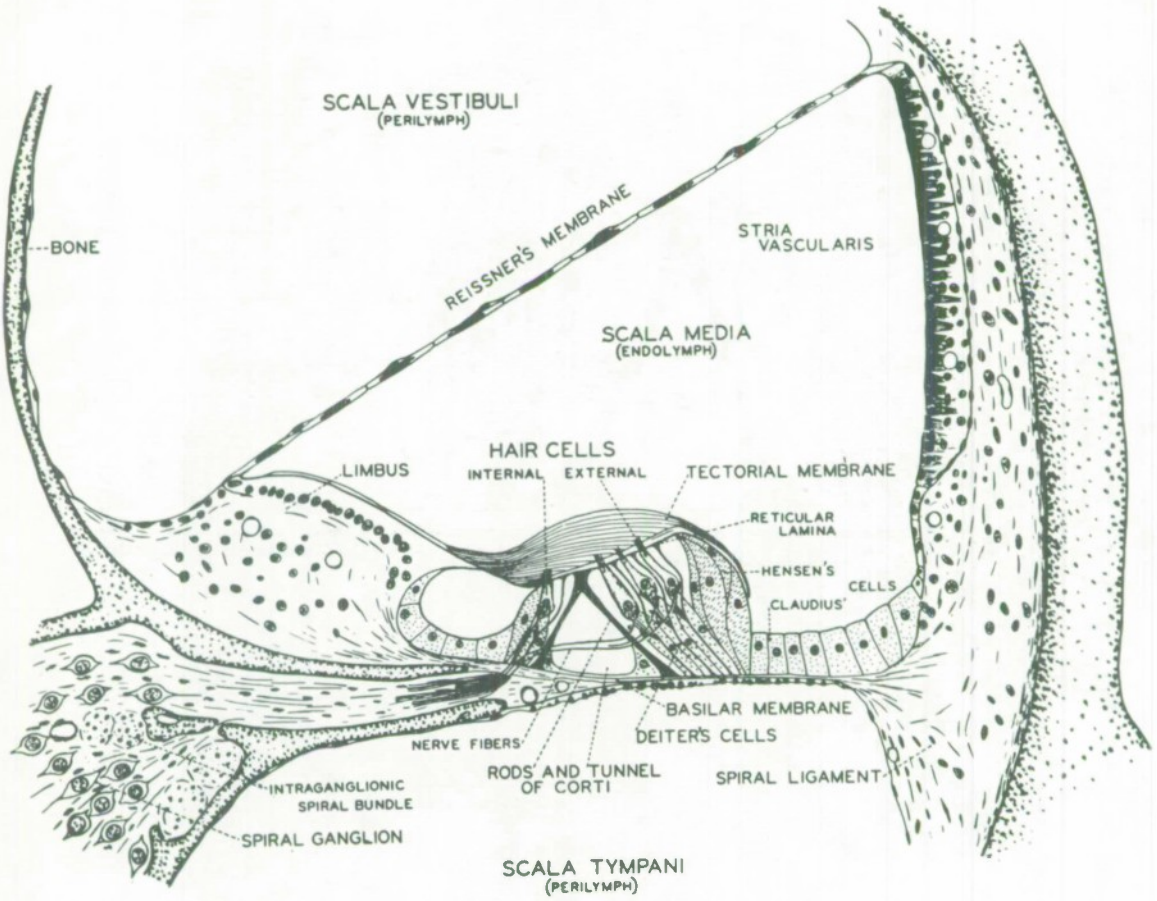

Fig. 2-4. Camera lucida drawing of a cross section of the cochlear partition in the second turn of a guinea pig cochlea. The attachment shown here of the tectorial membrane to the inner supporting cell, and to Hensen's cells, is based on microdissection of fresh, unfixed specimens. From Davis (1961).

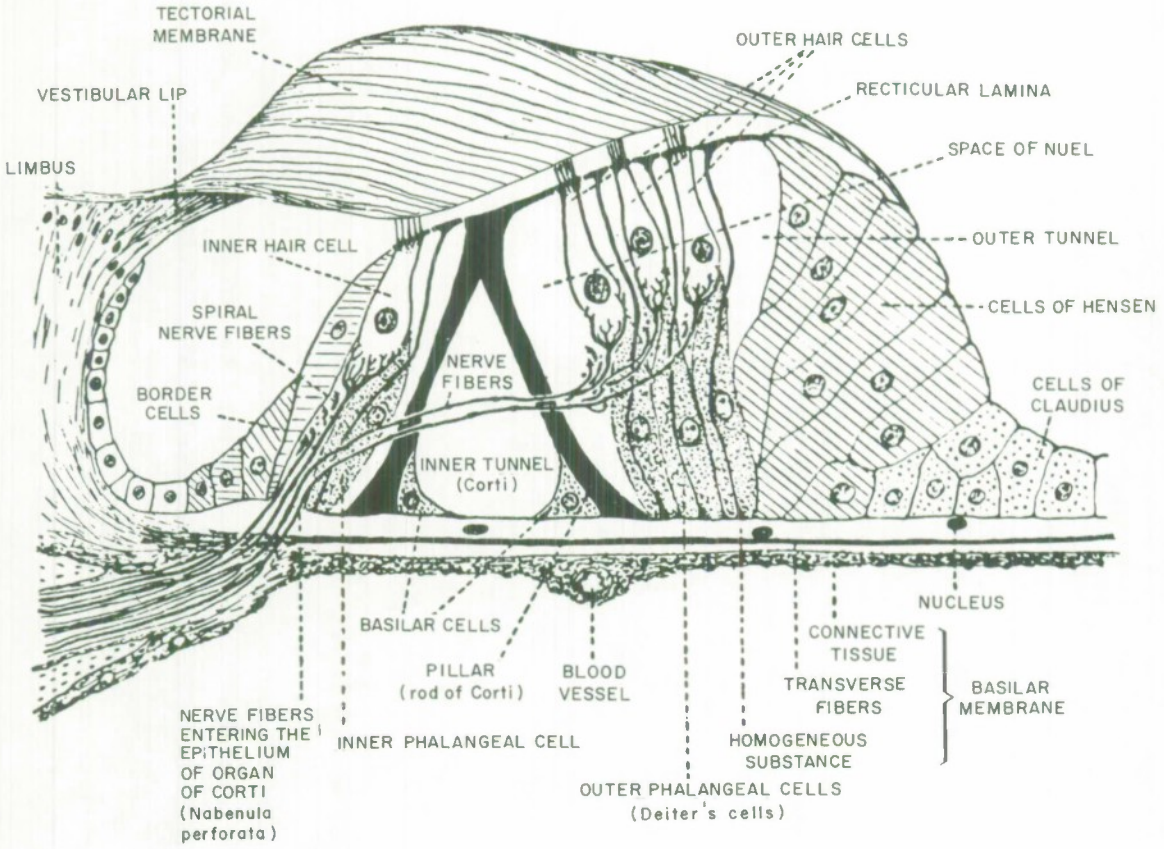

Fig. 2-5. Schematic diagram of the organ of Corti. From A. T. Rasmussen, Outlines of Neuroanatomy (modified) (1943).

It is these cells to which the mechanism of mechanical-to-neural transduction is attributed; it is these cells upon which the sensory endings of the nerve fibers of the primary neurons are situated. Therefore, it is in this vicinity that the auditory system is separated into mechanical and neural parts (see Fig. 2-1).

There are typically three rows of outer hair cells separated by the rods of Corti from a single row of inner hair cells. On the top of each hair cell is a tuft of hair which makes contact with the tectorial membrane. Figure 2-6 is an electron micrograph of a section of the cochlea taken parallel to the top of the organ of Corti; this section cuts through the hair tufts of about 10 hair cells. A remarkable, but not yet understood, feature is the uniform, W-like organization of the hairs on individual hair cells.

At the base of each hair cell are the endings of nerve fibers. Figure 2-7 is an electron micrograph of a section of the organ of Corti, but in the same plane as that of the sketch in Fig. 2-5, showing some of these endings. In addition, some hairs on top of the hair cells, as well as some nerve fibers crossing the tunnel of Corti, can be seen. Other nerve fibers also pass along the floor of the tunnel and continue on to make contact with hair cells. A majority of nerve fibers in the cochlea carry signals <u>from</u> hair cells to spiral ganglion cells--these are called afferent fibers. But some fibers carry signals <u>to</u> hair cells from some higher centers (as indicated by the feedback path shown in Fig. 2-1)--these are called efferent fibers.

Finally, on still closer examination (Fig. 2-8), it has been found that the nerve fiber endings on hair cells fall into two major categories: (1) small,

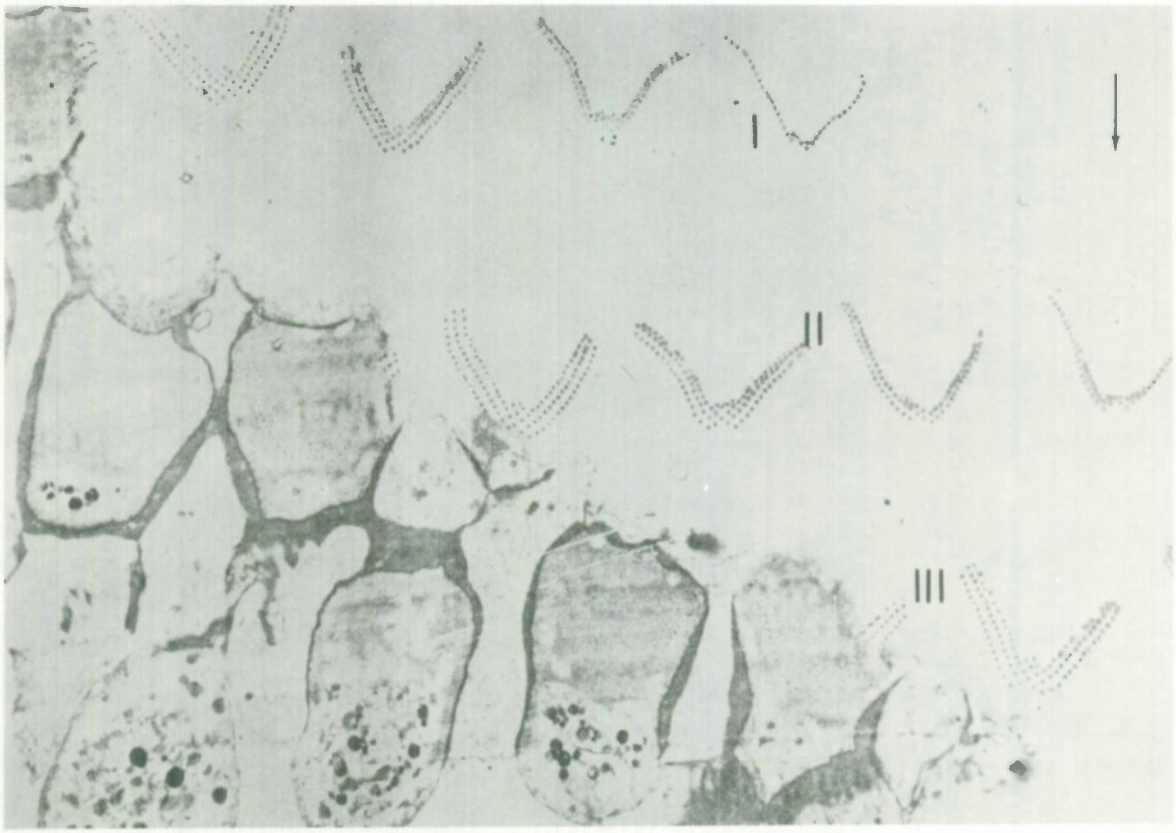

Fig. 2-6. The orientation of the sensory-hair bundles in the organ of Corti. The three rows of hair cells are marked I, II, and III. The arrow in the upper right corner points towards the Hensen's cells. Magnification X1800. From Flock et al. (1962).

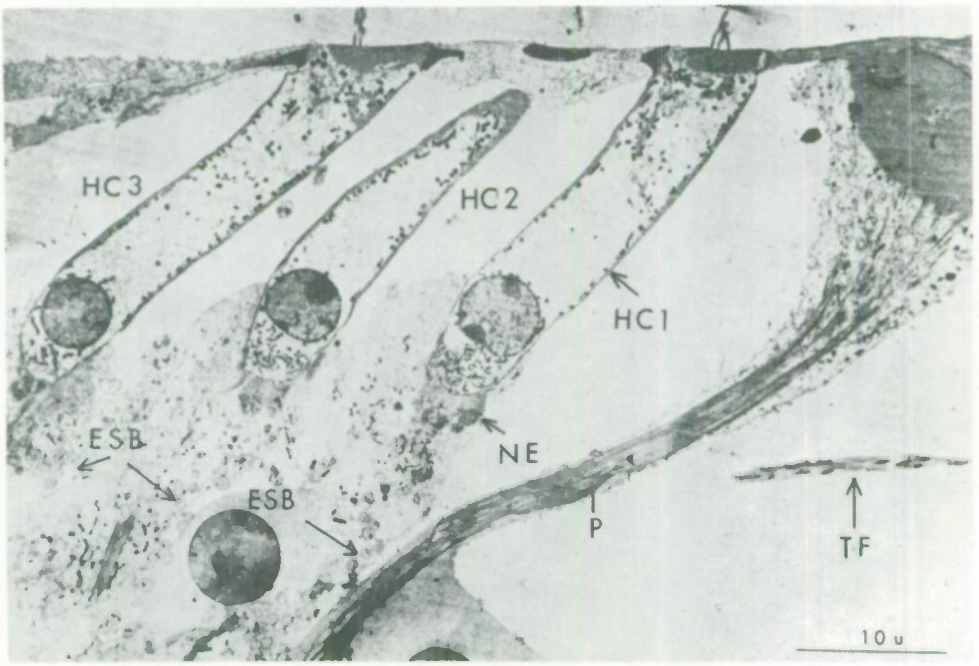

Fig. 2-7. Electron micrograph of normal organ of Corti showing hair cells, first row (HC), second row (HC2), third row (HC3), nerve endings (NC), tunnel fibres (TF), external spiral bundle (ESB), and outer pillar (P). Upper basal turn 9 mm area. From Kimura and Wersall (1963).

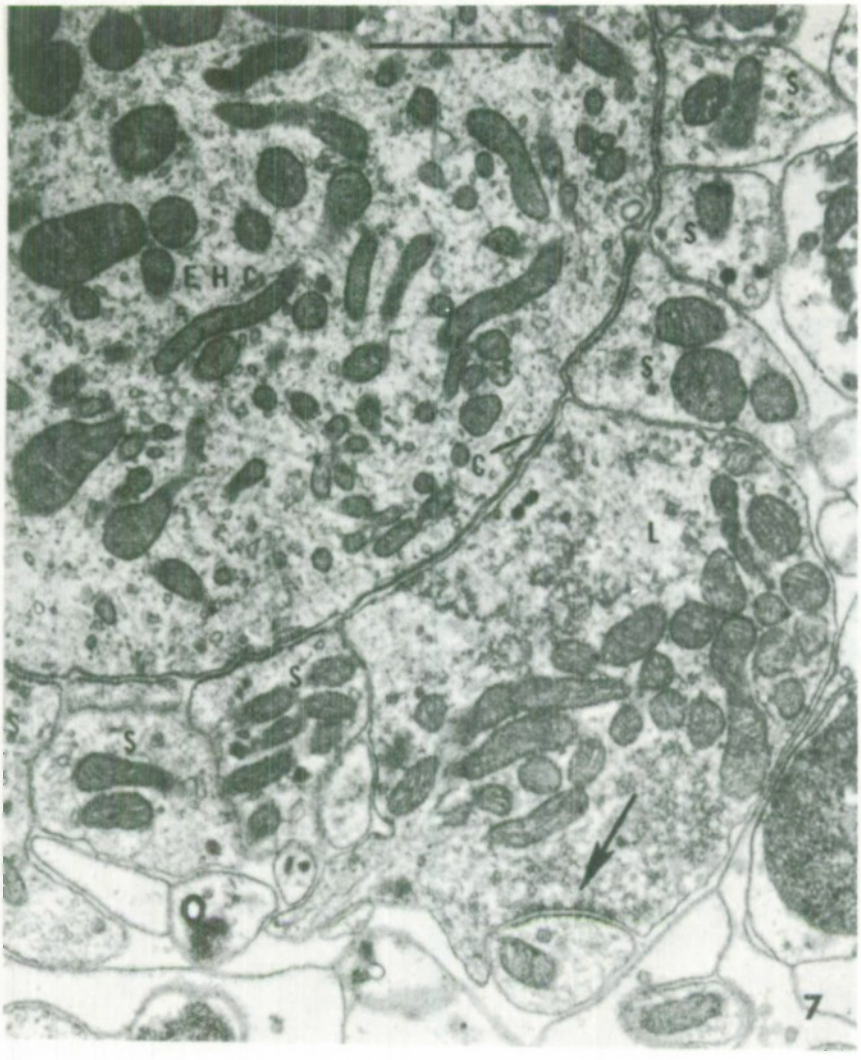

Fig. 2-8. Electron micrograph showing nerve endings on an external hair cell (EHC) of chinchilla. One large (L) nerve ending containing vesicles and mitochondria is present, with adjacent sub-synaptic cisterna (C) inside the hair cell. Several small nerve endings (S) are on either side. The large ending makes an apparent synaptic contact with another nerve (arrow) below. From Smith and Rasmussen (1963).

non-vesiculated endings (associated with afferents) and (2) large, vesiculated endings (associated with efferents). Vesicles are components of neural endings that presumably play a roll in the transmission of signals across the junctions between cell and ending. Thus those endings with vesicles are capable of transmitting a signal to the hair cell.

Returning to Fig. 2-6, it can be seen that the nerve fibers which make contact with the hair cells pass through the basilar membrane and enter the bony spiral lamina. These fibers proceed to the spiral ganglion (Fig. 2-5). The spiral ganglion is solely made up of bi-polar (one input fiber--one output fiber) neurons. These neurons are usually called primary neurons, so designated because they are the most peripheral neurons in the auditory system. The cochlear nerve, as seen in Fig. 2-4, is made up of the single-output fibers of these primary neurons. Within the cochlea, however, the input fiber to a primary neuron has many endings which may be connected to several hair cells; in turn, one hair cell may be connected to many primary neurons. Figure 2-9 shows a schematic representation of the course of the input fibers of the primary neurons, as well as the course of the efferent fibers entering the cochlea from higher centers. It is quite clear that the interrelations between a given primary neuron and different positions along the basilar membrane are quite complex and, it must be added, the details are not presently well known.

Higher Structures

We will not consider the structure of the higher centers in any detail. Nevertheless, in order to give an idea of the complexity, consider briefly

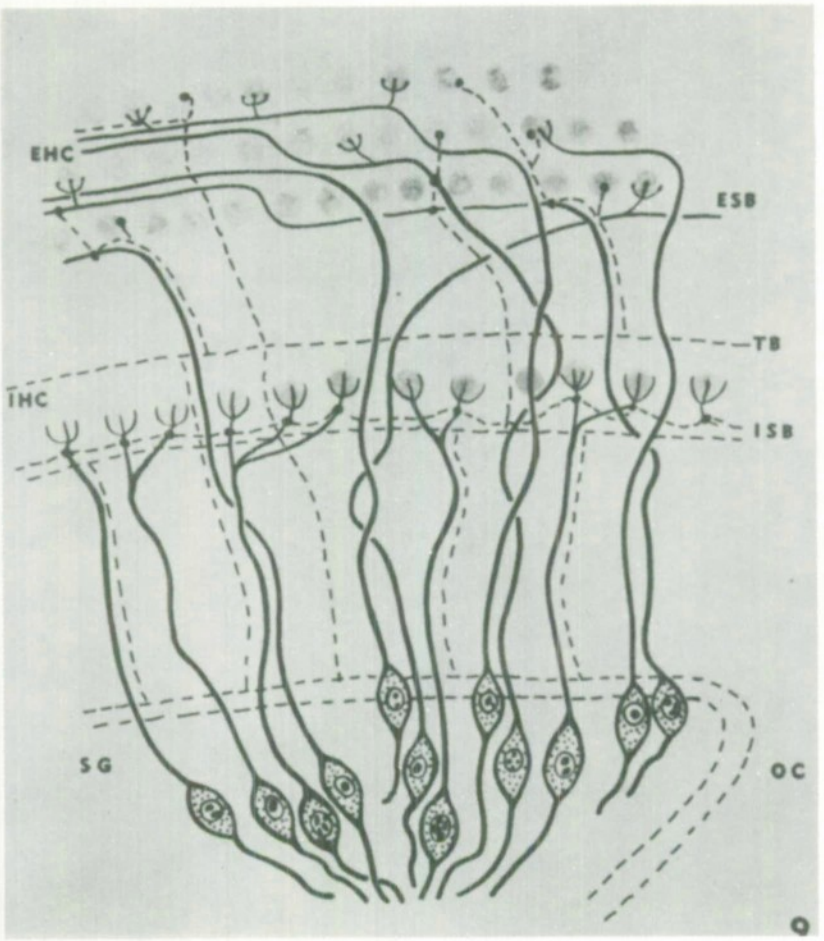

Fig. 2-9. Schematic diagram showing the course of the efferent olivo-cochlear nerve fibers in the organ of Corti of the chinchilla. The cochlear nerves are in solid black lines. The efferents are in dashed lines. EHC: external hair cells; ESB: external spiral bundles; ISB: internal spiral bundle; OC: olivo-cochlear bundle; SG: spiral ganglion; TB: tunnel bundle; IHC: internal hair cells. From Smith and Rasmussen (1963).

the destination of the fibers of the cochlear nerve. All of the cochlear nerve fibers terminate on cell bodies or dendrites within the cochlear nucleus. The cochlear nucleus has on the order of 100,000 neurons of which there are an estimated 40-50 different kinds based on physical appearance alone. It has been reported that each of the nerve fibers branches and has many endings, perhaps hundreds, on neurons corresponding to no less than 13 subdivisions of the cochlea nucleus. Similarly, some neurons in the cochlear nucleus receive endings from a few, while others perhaps hundreds of, cochlear nerve fibers, in addition to endings from other cells within the nucleus and endings from cells located in higher structures (efferent inputs). Clearly, the interconnections and relationships--that one might consider to be the wiring diagram--are extremely complex. Although they are obviously contributors and determinors of behavior in perception of speech and sound, knowledge of their anatomical details is not necessary for our preliminary objectives.

Summary

In this brief sketch, we have seen the basic make-up of that portion of the auditory system which is essentially peripheral to the brain stem. Although we limit ourselves to this region, we will subdivide it into two major parts--mechanical and non-mechanical. As indicated earlier, the mechanical part relates to the sound-transmitting structures and the non-mechanical part relates to the neural structures. This division is motivated for the following reasons:

1. The mechanical subsystems are linear in the dynamic range of interest to us and, hence, they can be conveniently characterized by

impulse responses or frequency responses. There exists enough reasonably accurate data on the mechanical subsystems to specify the linear models with some confidence.

2. The neural subsystems are non-linear and, hence, their characterization is more complicated and requires more data than that of linear systems.

Anatomy and Physiology References

J. B. Deaver, Surgical Anatomy of the Human Body (The Blakiston Company, 1926).

H. Davis, "Peripheral Coding of Auditory Information," Sensory Communications, edited by W. A. Rosenblatt, 119-141 (M. I. T. Press, 1961).

H. Davis, "Some Principles of Sensory Receptor Action," Physiological Review 41, 391-416 (1961).

A. Flock, R. Kimura, P. Lundquist, J. Wersäll, "Morphological Basis of Directional Sensitivity of the Outer Hair Cells in the Organ of Corti," J. Acoust. Soc. Am. 34, Part 2, 1351-1355 (1962).

R. Kimura, J. Wersäll, "Termination of the Olivo-Cochlear Bundle in Relation to the Outer Hair Cells of the Organ of Corti in Guinea Pig," Acta Otolaryng. 55, 11-32 (1963).

Y. Nomura and H. F. Schuknecht, "The Efferent Fibers in the Cochlea," Ann. Otol. Rhinol. and Laryngol. 74, 289-303 (1965).

A. T. Rasmussen, Outlines of Neuroanatomy, Third Edition (William C. Brown Co., Dubuque, Iowa, 1943).

C. A. Smith, G. L. Rasmussen, "Recent Observations on the Olivo-Cochlear Bundle," Ann. Otol. Rhinol. and Laryngol. 72, 489-507 (1963).

H. H. Spoendlin and R. R. Gacek, "Electron-Microscopic Study of the Efferent and Afferent Innervation of the Organ of Corti in the Cat," Ann. Otol Rhinol. and Laryngol. 72, 660-686 (1963).

B. von Bekesy, Experiments in Hearing, edited by E. G. Wever (McGraw-Hill, Inc., New York, 1960).

F. M. Weiner, "How We Hear," Physics Today 2, 8-14 (1949).

E. G. Wever, M. Lawrence, Physiological Acoustics (Princeton University Press, Princeton, New Jersey, 1954).

"Neurophysiology," Handbook of Physiology, Vol. 1, Sect. 1, Am. Phys. Soc. (Washington, D. C., 1959).

CHAPTER III

Mechanical (Sound-Transmitting) Portion of Auditory System

A. THE EXTERNAL EAR

Experimental observations of the effects on sound transmission that are introduced by the pinna and auditory canal indicate that sound pressure levels at the eardrum relative to free-field pressures* are dependent on the orientation of the source (due to diffraction by the head and pinna), while sound pressures at the eardrum relative to those at the entrance of the auditory canal are not. Presently we will ignore the diffraction effects of the head and will only consider the effects of the ear canal--as, for example, is the case for performance with earphones.

The auditory canal acts as an effective pressure "transformer" over a broad range of frequencies. This can be seen from the plot of the ratio in dB of the magnitudes of sound pressure at the eardrum to sound pressure at the entrance of the auditory canal (Fig. 3-1). The maximum gain of about 11 dB is realized near 4 kHz. Since the auditory canal is an approximately straight tube, it is not unreasonable to expect it to have acoustical properties similar to a rigid walled cylinder. Weiner and Ross modeled the canal as a rigid-walled cylinder of 2.3 cm in length, with a rigid termination. Cross-sectional dimensions were assumed small compared to an acoustical wavelength. The results of their model are also shown in Fig. 3-1. This model,

---

* Free-field pressures are usually those which are measured at a point near the center of the head in absence of the subject.

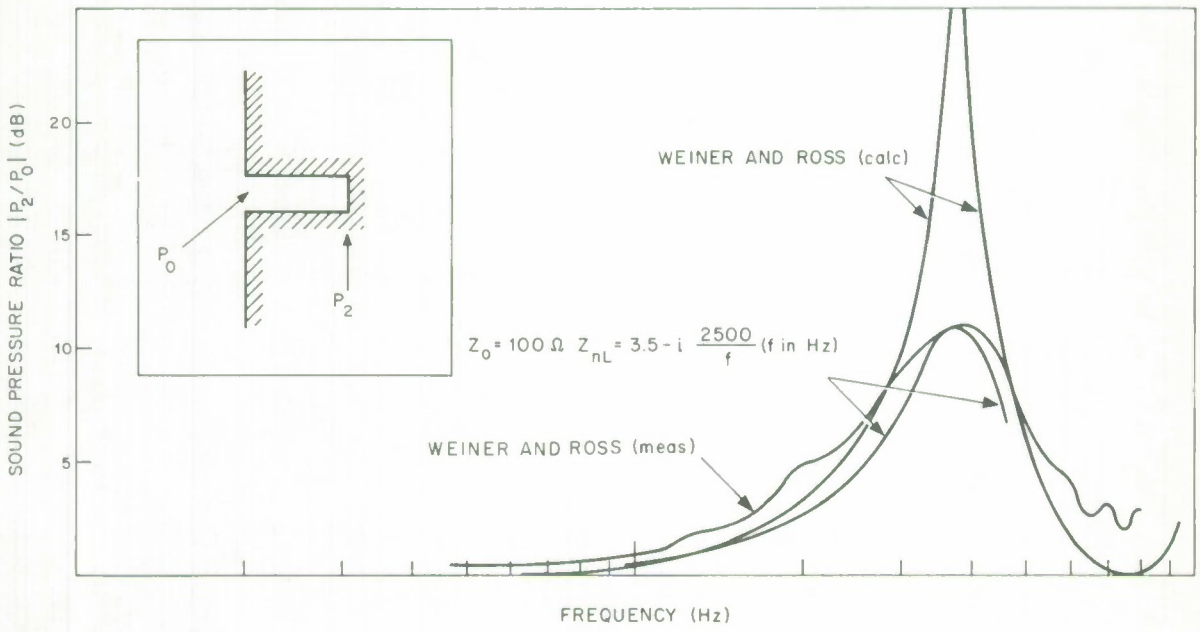

Fig. 3-1. Sound transmission characteristics of the external ear.

of course, is a lossless acoustic transmission line and will thus present an infinite pressure ratio at frequencies corresponding to odd multiples of the quarter-wavelength ($\lambda/4$) resonant frequency.

Refining this idealized situation, we have assumed as a model of the external ear canal a rigid tube (also 2.3 cm in length) but with a flexible termination having an acoustic impedance corresponding to that of the eardrum loaded by the middle ear. The measured acoustic input impedance of the middle ear as seen from the outside of the eardrum is shown in Fig. 3-2. We, therefore, take as the termination of our transmission line model of the ear canal the impedance

$$Z_L = 350 - i\frac{2.5 \times 10^5}{f} \text{ ohms (f in Hz)}$$

which is a simple approximation to the data of Fig. 3-2 which is valid up to a few kHz. We take the characteristic impedance of our transmission line model to be 100 ohms (Ω) [corresponding to an ear canal cross-sectional area of 0.414 cm^2, approximately the average area in humans). The pressure ratios obtained from this model are also shown in Fig. 3-1. Our loaded transmission line model provides a pressure ratio which matches the measured data quite adequately for our purposes.

Summary

A satisfactory electrical analog for the external ear would be an electrical transmission line with a characteristic impedance of 100 Ω which is a

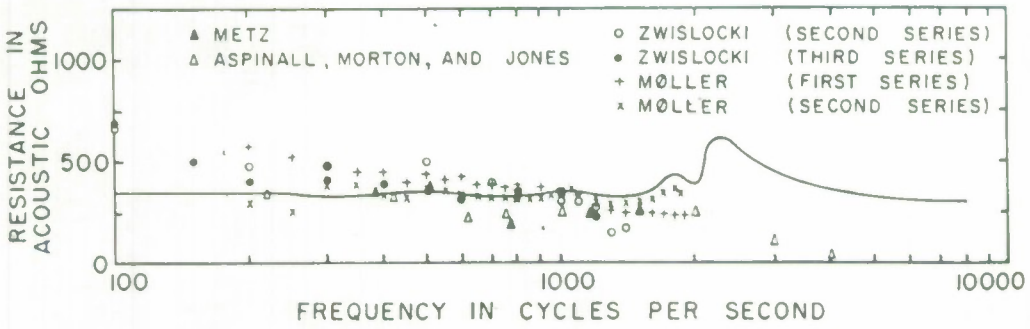

Fig. 3-2(a). Acoustic resistance at the eardrum. The points indicate the average values obtained in several experimental series. From Zwislocki (1965).

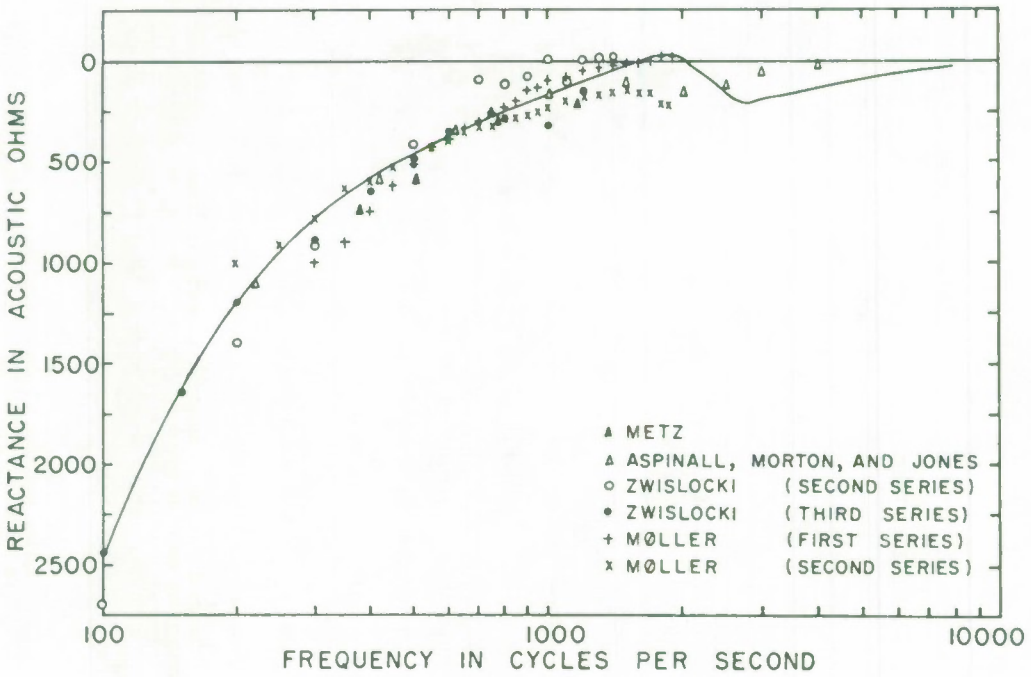

Fig. 3-2(b). Acoustic reactance at the eardrum. The points indicate the average values obtained in several experimental series; the curve was obtained on the network analog. From Zwislocki (1965).

quarter wavelength long at 3.7 kHz,* and with a normalized termination or load impedance of

$$Z_{nL} = \frac{Z_L}{Z_0} = 3.5 - i\,\frac{2500}{f}. \qquad (3\text{-}1)$$

* An acoustic plane wave in air of frequency 3.7 kHz has a quarter wavelength of 2.3 cm.

B. THE MIDDLE EAR

The middle ear mechanism, consisting of the tympanic membrane, the ossicles, and the middle ear cavities, provides an acoustic coupling between the external ear canal and the fluid of the cochlea (inner ear). Pressure variations in the external ear canal corresponding to a sound wave cause flexing of the eardrum which results in movement of the ossicles. Displacement of the stapes footplate, which acts like a driving piston in the wall of the cochlea, launches an acoustic wave in the cochlear fluid (see Fig. 3-3). In studying the middle ear, the cochlear is the "load" driven by the ossicles and must, therefore, be considered here too. However, the detailed mechanics of the cochlea itself are discussed in the following section.

Clearly, an adequate characterization of the middle ear behavior would be the transformation expressing stapes velocity $u_s(t)$ [or displacement $y_s(t)$] as a function of time t in terms of the excess pressure in the ear canal at the drum, $p_d(t)$, as a function of time. Several investigators have studied this transformation experimentally in human cadavers as well as in anesthetized cats and guinea pigs. They have established that this transformation can be considered linear at least for a range of sound intensities below the threshold of pain and a frequency range up to about 10 kHz. The upper limits on intensity corresponding to linear operation are well known but the lower limits are not because measurements are difficult to make at very low intensities. Nevertheless, in the intensity and frequency range of speech sounds, the middle ear mechanism can safely be taken as a linear one. Of course, this greatly eases the problem of compactly characterizing the behavior of the middle ear.

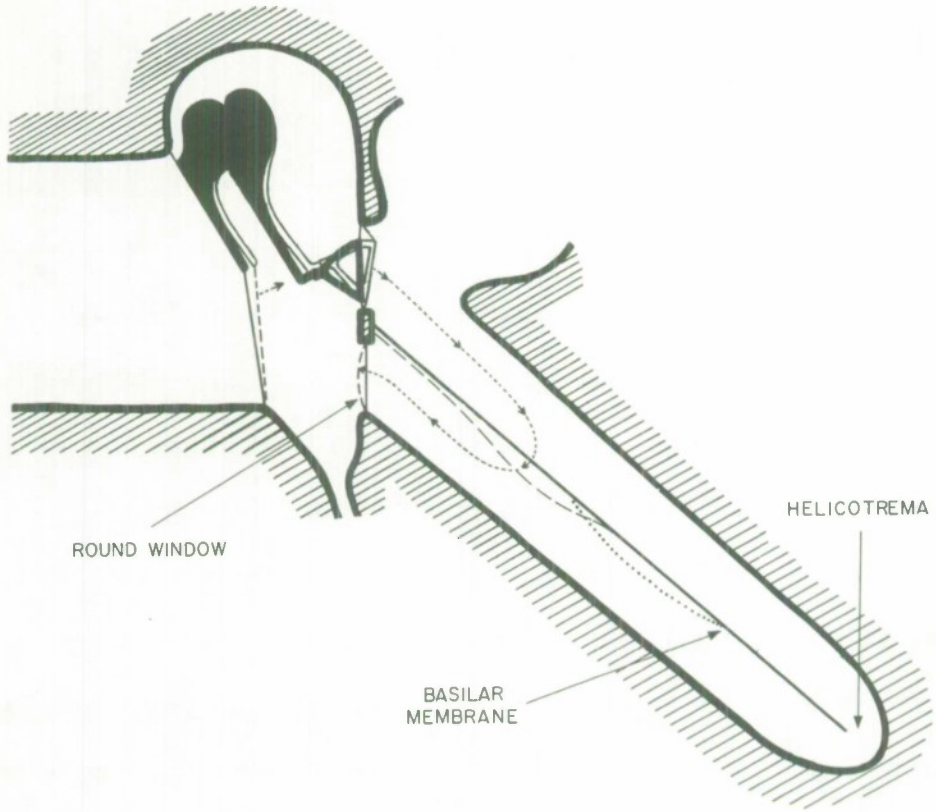

Fig. 3-3. Schematic diagram of the tympanic membrane, the ossicles, and the basilar membrane. The solid figures of the ossicles and the solid lines for the tympanic, the basilar, and the round-window membranes show the positions of these structures at rest. The open outlines of the ossicles and the broken lines for the membranes show their positions following inward displacement of the tympanic membrane by a sound wave. The dotted line shows the disturbance traveling along the basilar membrane at a later instant. After Stevens and Davis (1938).

In its linear range, the middle ear can be characterized by the transfer function relating the amplitude and phase of stapes displacement to that of the driving pressure $p_d(t)$ for sinusoidal stimuli. In Fig. 3-4 we show the direct measurements of the middle ear transfer function made by von Bekesy on two human temporal bones. The plot shows relative stapes amplitude for a constant level of sinusoidal sound pressure at the eardrum. These measurements were made by capacitive probe techniques. There is obviously a significant difference between these two transfer functions, even within the restricted frequency range covered. In Fig. 3-5 we show a sample of the corresponding measurements recently made by Peake and Guinan on anesthetized cats. These latter measurements were made through a microscope under stroboscopic illumination. The general appearance of all the data shows the middle ear to be a low-pass device with a natural frequency of approximately 1000 Hz. But the transfer functions obtained from human temporal bones exhibit a more damped resonance than that for the anesthetized cats. The data of Peake and Guinan exhibits less variability and covers a much broader frequency range. We will, therefore, rely on this data in our modeling of the middle ear even though it is cat data.

The eardrum is not under very much tension and has a high frequency resonance (at about 10 kHz). At frequencies below 4 kHz, therefore, the eardrum may be considered a light but rigid piston with a restoring spring. We also assume that the motion of the handle (manubrium) of the malleus is exactly the same as the drum motion in this frequency range. The eardrum, in addition to driving the ossicles, also acts as the driving piston for the

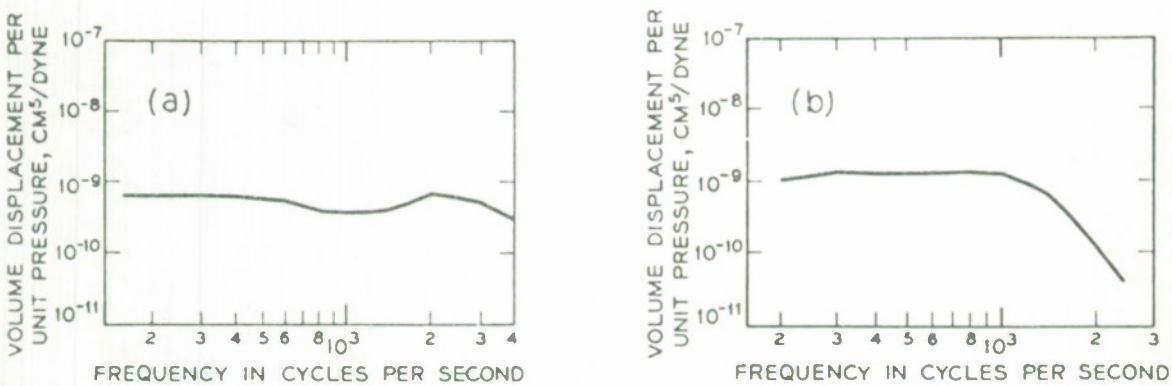

Fig. 3-4. Data on middle ear transmission; effective stapes displacement for a constant sound pressure at the eardrum. (a) von Bekesy, 1960 (one determination); (b) von Bekesy, 1960 (another determination). From Flanagan.

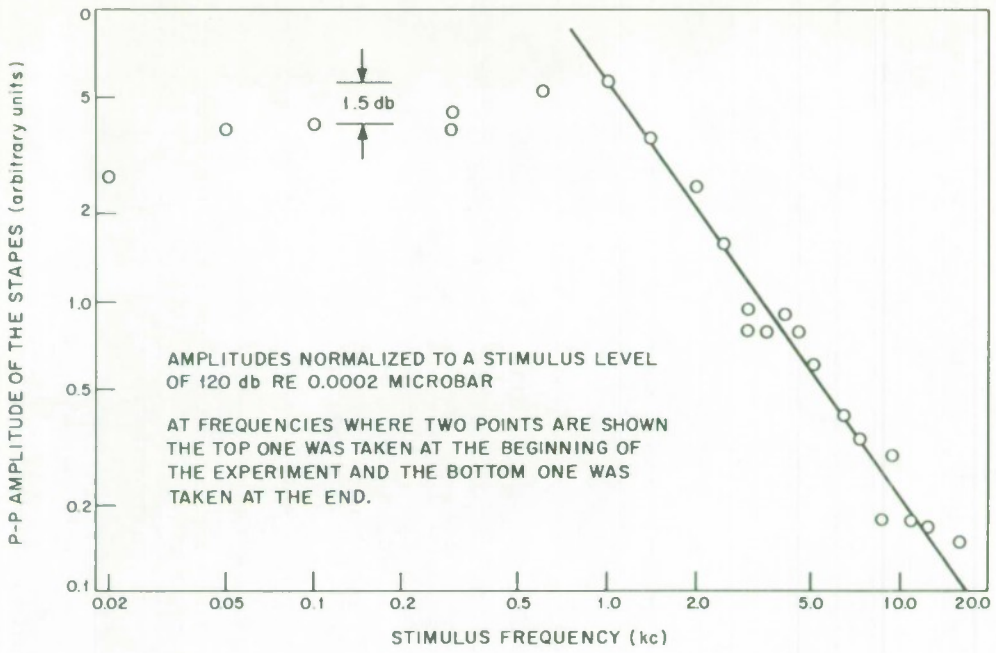

Fig. 3-5(a). Stapes' amplitude as a function of frequency in Cat 30. From Guinan.

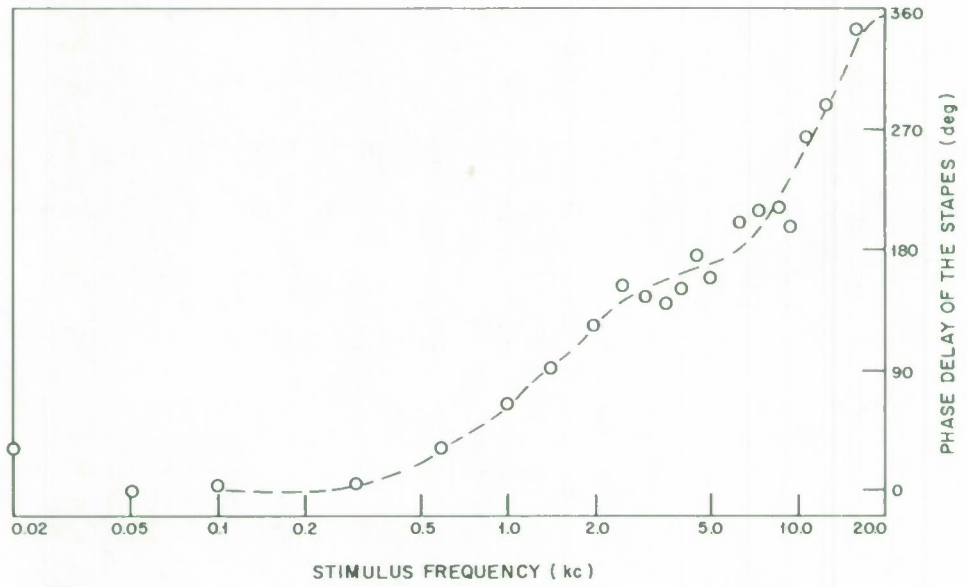

Fig. 3-5(b). Phase delay of the stapes' amplitude after stimulus pressure in Cat 30. From Guinan.

air-filled, middle ear cavities. A mechanical schematic of the middle ear is shown in Fig. 3-6. The block diagram in Fig. 3-7a illustrates the mechanical interconnections of the various parts of the middle ear.

Middle Ear Circuit Model

In constructing an electrical circuit analog of the middle ear, we take electrical current as the analog of mechanical velocity and voltage as the analog of force. From Fig. 3-7a, we see the velocity of the eardrum, the handle of the malleus, and the air in the cavities are one and the same (under our assumption about the eardrum), while the driving pressure on the eardrum is resisted by the sum of the forces due to the drum elasticity, the handle, and the air of the cavities. Therefore, the circuits representing these parts of the middle ear are in series in the circuit model.

Due to elasticity between the ossicles, the stapes will not move exactly in phase with the handle of the malleus. Furthermore, since the ossicles are connected to the temporal bone by supporting structures, forces are transmitted to the temporal bone. Thus, the force on the handle of the malleus (as well as the velocity) is not directed entirely to the stapes. Therefore, a general series-parallel network is needed for the ossicles in the circuit model. A block diagram of the electrical analog of the middle ear is shown in Fig. 3-7b.

The eardrum and ossicles closely resemble a simple mechanical system of pistons, levers, springs, and dashpots. Such a system can be modeled by the lumped electrical circuit analog illustrated in Fig. 3-8a. We have represented the cochlear load in the circuit model as a resistor since this is the predominant feature of the input impedance of the cochlea.

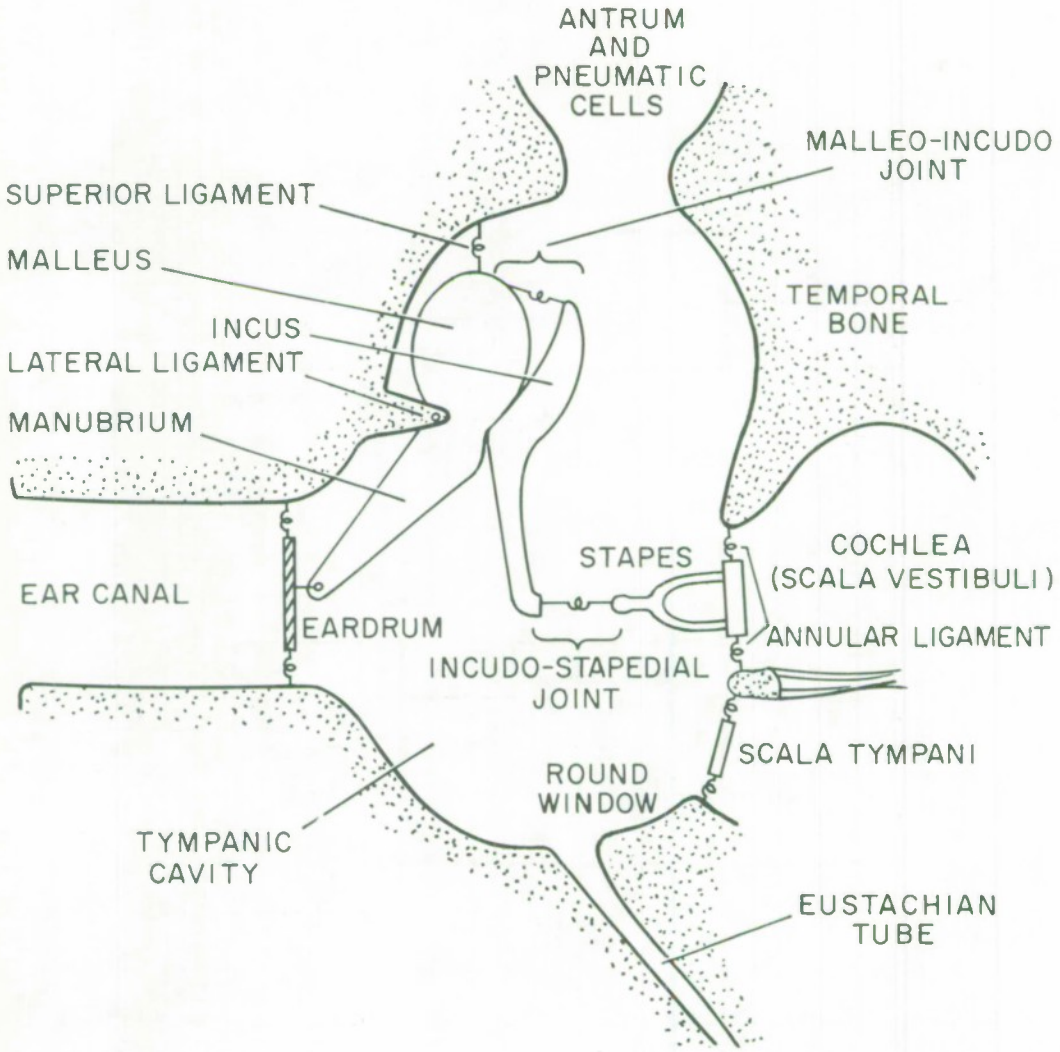

Fig. 3-6. Mechanical schematic of the external and middle ear.

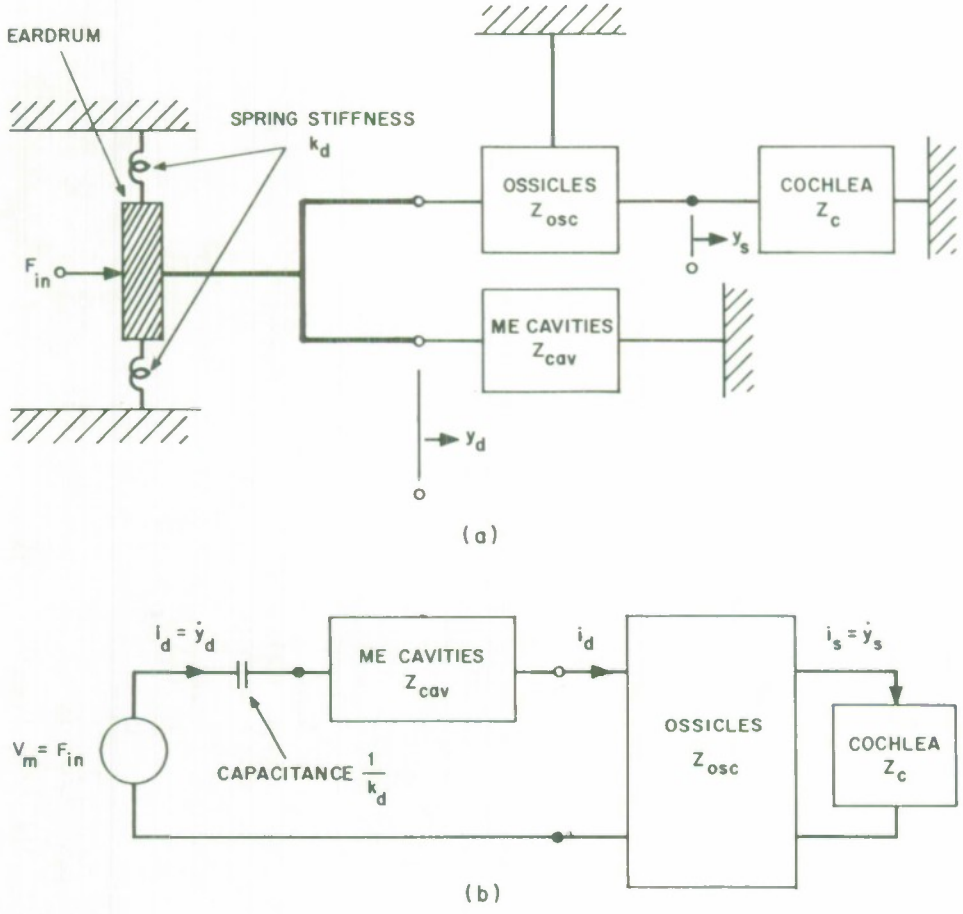

Fig. 3-7. Block diagram of middle ear models: (a) mechanical model; (b) electrical analog.

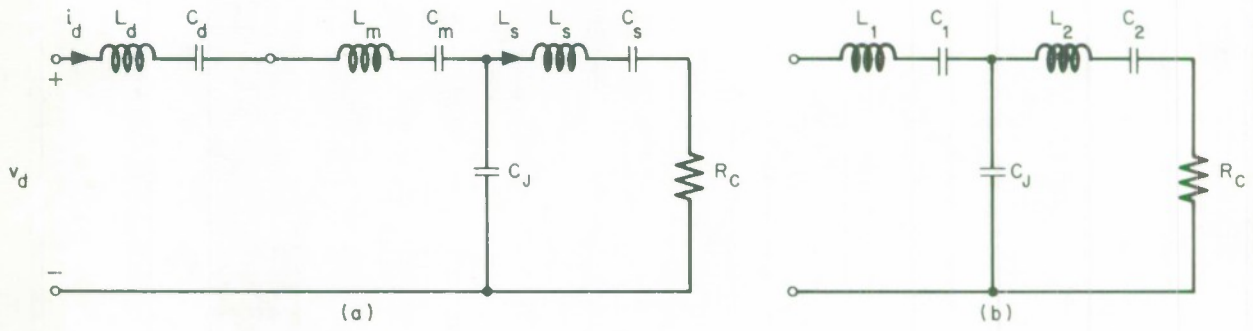

Fig. 3-8. Electrical circuit analog of eardrum and ossicles.

In our schematics, we have shown the malleo-incudo joint to be elastic. The incudo-stapedial joint is most frequently taken as being the elastic joint but recent experimental evidence seems to indicate the malleo-incudo joint may be the more elastic. The capacitance C_J in the circuit can represent the elasticity of either of the joints by simply including the incus parameters in the malleus parameters, L_m and C_m, if the incudo-stapedial joint is taken to be elastic or by including the incus parameters in the stapes parameters, L_s and C_s, if the malleo-incudo joint is taken as the elastic one. The circuit model in Fig. 3-8b is general enough for either of these cases. We take the malleo-incudo joint as the elastic one, in agreement with Peake and Guinan.

Peake and Guinan (1965) report that for cat the stapes and malleus are in phase at least up to 5 kHz, which indicates the joint stiffness, K_J, is large enough to be considered a rigid connection in this frequency range. Thus C_J can be considered to be an open circuit and the middle ear circuit model becomes a simple RLC circuit in this frequency range. Assuming the pressure at the eardrum $p_d(t)$ to be sinusoidal,*

$$p_d(t) = \mathrm{Re} P_d(s)\, e^{st} , \quad s = i2\pi f ,$$

corresponding to the voltage drive† of the middle ear circuit, the current in the cochlear resistor corresponds to the stapes velocity $u_s(t) = \mathrm{Re} U_s(s)\, e^{st}$ and the circuit transfer function is

---

* "Real part of" is abbreviated as Re.

† We do not have a quantitative relation between the mechanical system and analog circuit variables, so we can loosely take a voltage as the analog of a pressure rather than a force.

$$\frac{U_s}{P_d} \approx \frac{1}{sL + R_c + \frac{1}{Cs}}, \quad |s| \leq 2\pi \cdot 5000$$

where

$$L = L_1 + L_2, \quad C = \frac{C_1 C_2}{C_1 + C_2}$$

The transfer function relating stapes displacement,

$$y_s(t) = Y_s e^{st} = \frac{U_s}{s} e^{st}$$

is then

$$\frac{Y_s}{P_d} = H_{me}(s) \approx \frac{1}{L\left(s^2 + \frac{R_c}{L} s + \frac{1}{LC}\right)}, \quad |s| \leq 2\pi \cdot 5000$$

If we choose the parameters

$$\frac{R_c}{L} = 2\pi \cdot 2110$$

$$\frac{1}{LC} = (2\pi \cdot 1500)^2$$

to match the main middle ear resonance at 1000 Hz, this simplified circuit model gives the transfer function [normalized to $H_{me}(0)$] shown in Fig. 3-9. This model is obviously unsatisfactory since both the rolloff or slope (in dB/octave) of $H_{me}(s)$ above the resonance and the phase characteristics are not even similar to that of the measured data.

Taking the incudo-stapedial joint as elastic, we reinsert C_J in the circuit model; but neglecting L_2 and C_2 (stapes mass and annular ligament stiffness), the transfer function now becomes,

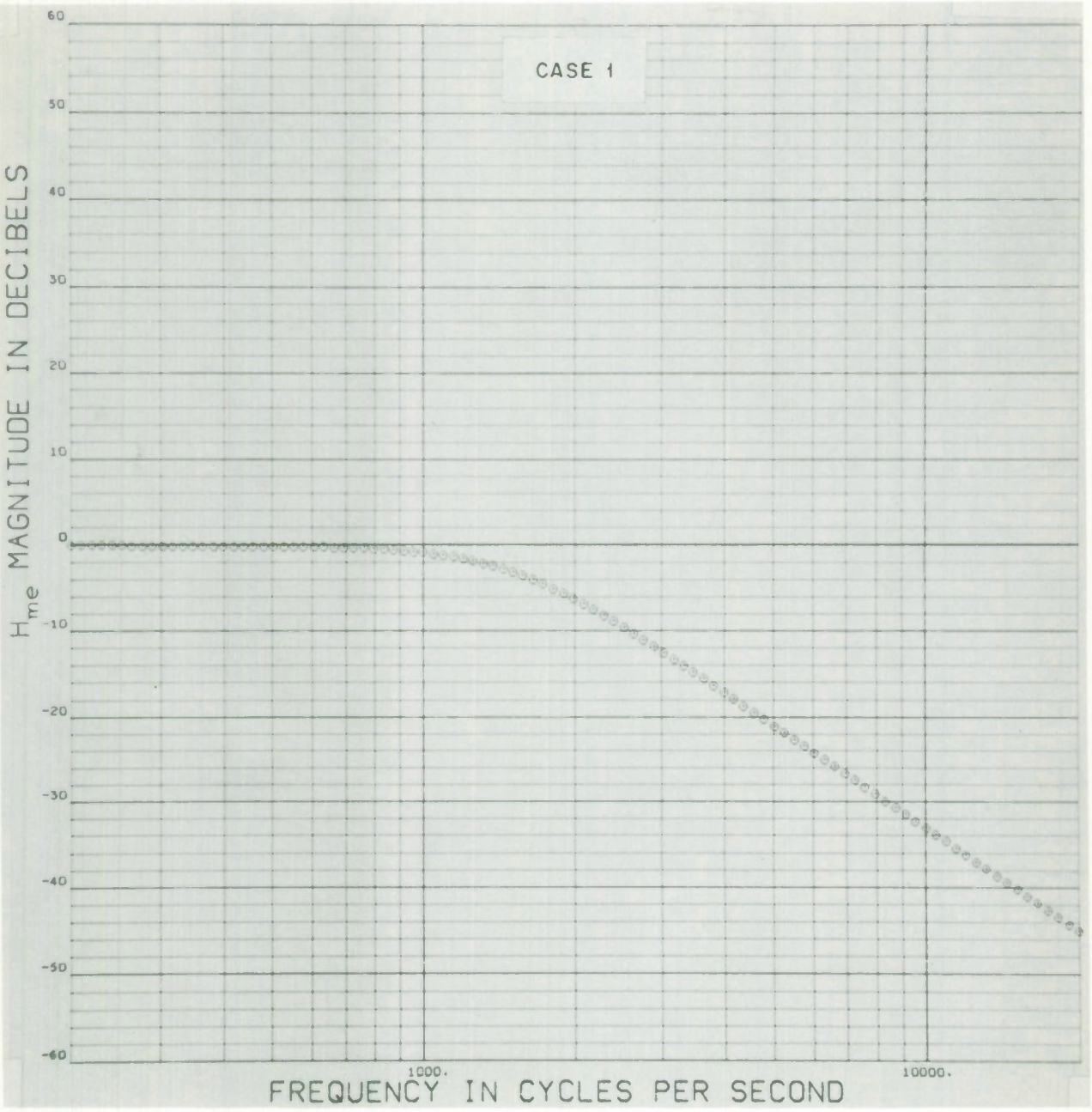

Fig. 3-9. Transfer function $Y_s(s)/P_d(s)$ for simple middle ear circuit analog (C_J open circuit).

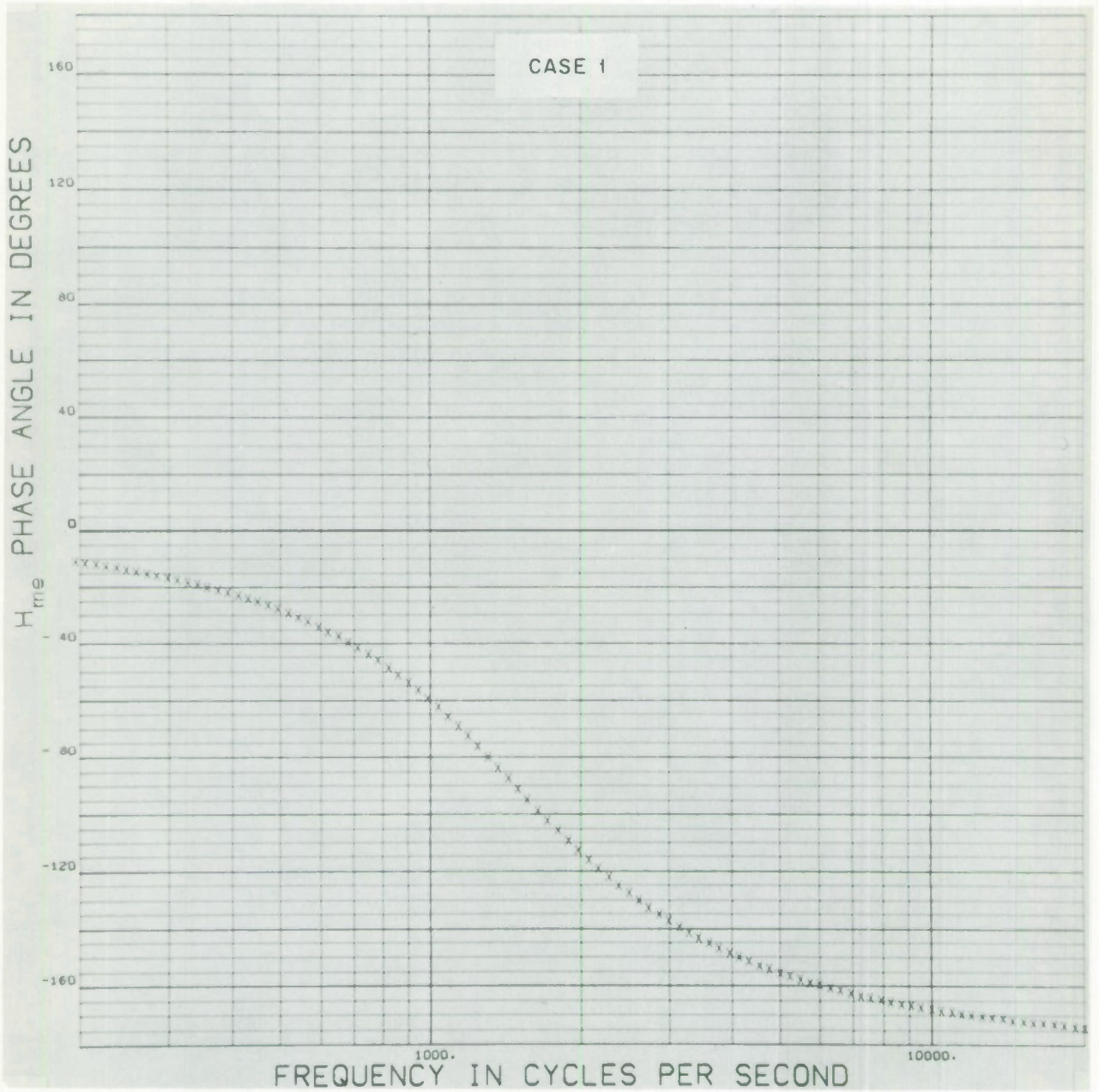

Fig. 3-9. (Continued).

$$H_{me}(s) = L_1[(s^2 + \frac{1}{L_1C_1})(RC_Js + 1) + \frac{R}{L_1}\, s] \qquad (3\text{-}2)$$

If we take

$$\frac{1}{RC_J} = 2\pi 9800$$

corresponding to the fact that the malleus and stapes are in phase at least up to 5000 Hz, we get the normalized transfer function shown in Fig. 3-10. This results in an increase in the rolloff above 1000 Hz, but the phase characteristic now approaches an asymptote of -270° for very high frequencies (well above 10 kHz) which is still not in agreement with the measured phase characteristic.

The transfer function corresponding to the general circuit of Fig. 3-8b is

$$H_{me}(s) = \frac{Y_s}{P_d} = \frac{1}{L_1[(s^2 + \frac{1}{L_1C_1})(L_2C_Js + RC_Js + 1 + \frac{C_J}{C_2}) + (\frac{L_2}{L_1}s^2 + \frac{R}{L_1}\, s + \frac{1}{L_1C_2})]} \qquad (3\text{-}3)$$

This can be expressed as

$$H_{me}(s) = \frac{1}{L_1[(s^2 + \omega_1^2)(\frac{\rho}{2\alpha_1\alpha_J}\, s^2 + \frac{s}{\alpha_J} + 1 + \frac{\omega_1^2\sigma}{2\alpha_1\alpha_J}) + (\rho s^2 + 2\alpha_1 s + \sigma\omega_1^2)]} \qquad (3\text{-}4)$$

where

$$\rho = \frac{L_2}{L_1}$$

$$\sigma = \frac{C_1}{C_2}$$

and the definitions of the parameters ω_1^2, α_1, α_J are clear from the previous equation. W. T. Peake* has obtained a good fit to his measurements of $H_{me}(s)$ by choosing $\rho = 1/3$ and keeping the total inductance $L_1 + L_2$ the same as in the previous cases (with $C_2 \to \infty$, or $\sigma = 0$). This is shown in Fig. 3-11. The rolloff and phase characteristic obviously match the measurements very well. If we test the sensitivity of the model by now dividing the capacitance C_1 of this circuit equally between C_1 and C_2, i.e., setting $\sigma = 1$ but keeping $C_1C_2/(C_1+C_2)$ the same as the previous C_1, we get the transfer function shown in Fig. 3-12, which is almost indistinguishable from the previous case. This insensitivity implies that it would be impossible to deduce which of these last two circuits actually applies because of the level of precision of the measured transfer functions.

It is difficult to make any direct correspondence between circuit element parameters and mechanical system parameters. For instance, the mass of the malleus does not translate directly into a circuit inductance value because the motion of the malleus is a pivotal one, not purely a translation. Therefore, lever arms and moments of inertia are involved in determining the circuit inductance corresponding to the malleus. Damping constants associated with specific joints of the mechanical system cannot be measured statically. Spring constants can be determined from static measurements in some cases. The insensitivity of the circuit model to changes in certain elements makes possible only a loose association of these element values with mechanical properties of the parts of the middle ear. Therefore, it is very risky to

---

* We are indebted to Peake for his circuit element values for all of the models for $H_{me}(s)$.

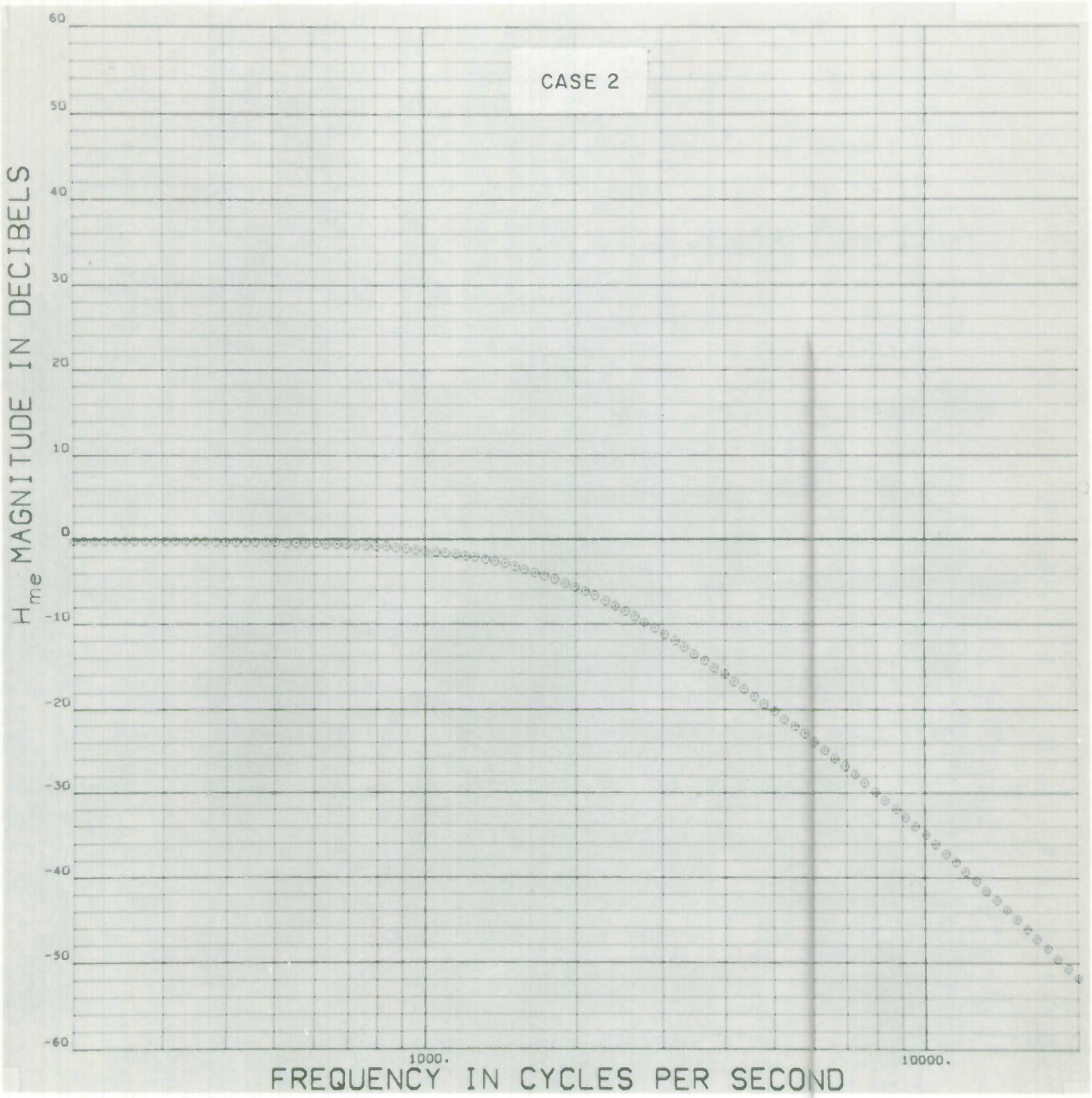

Fig. 3-10. Transfer function of middle ear circuit analog with elastic incudo-stapedial joint (L_2 and C_2 neglected).

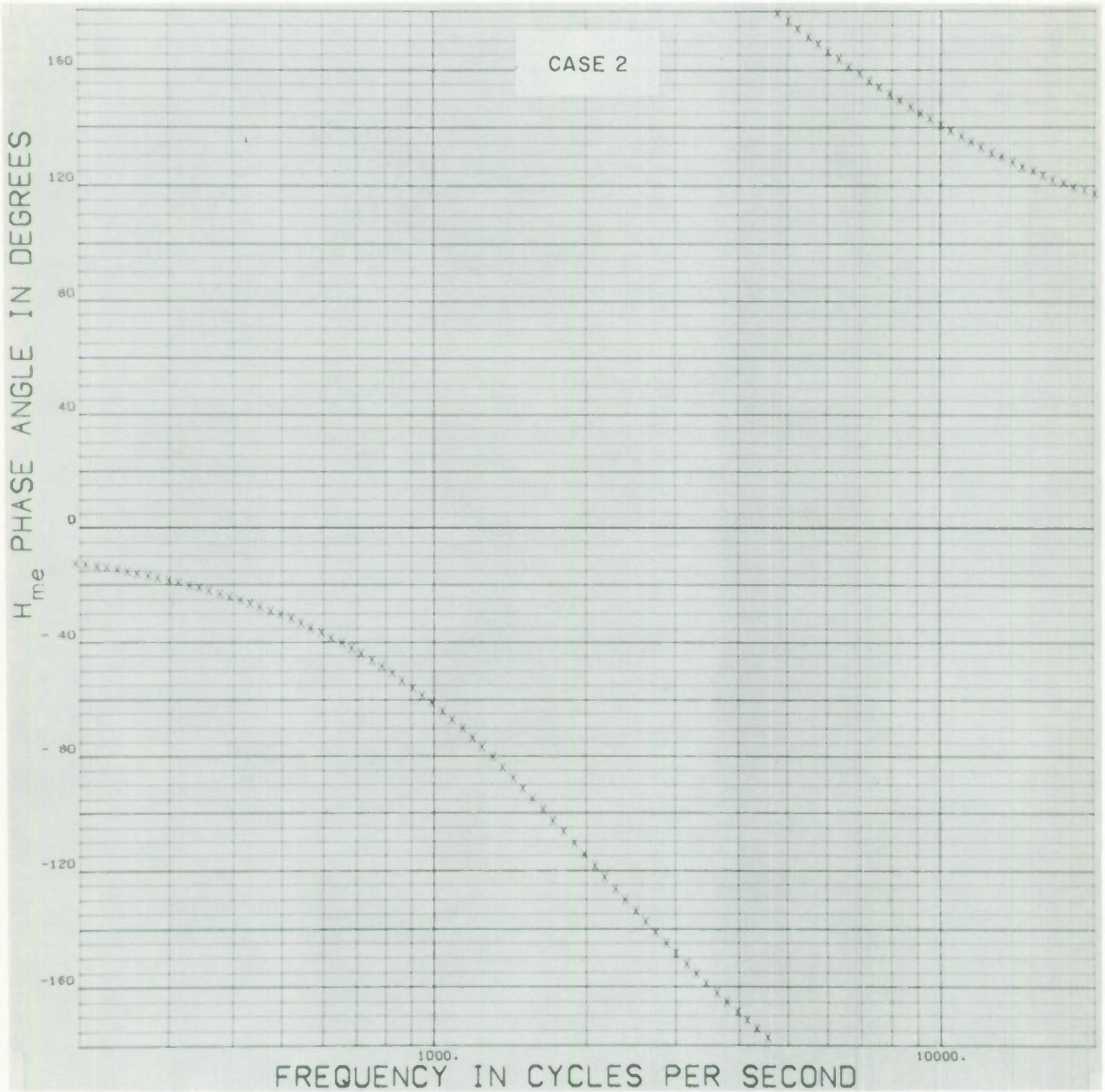

Fig. 3-10. (Continued).

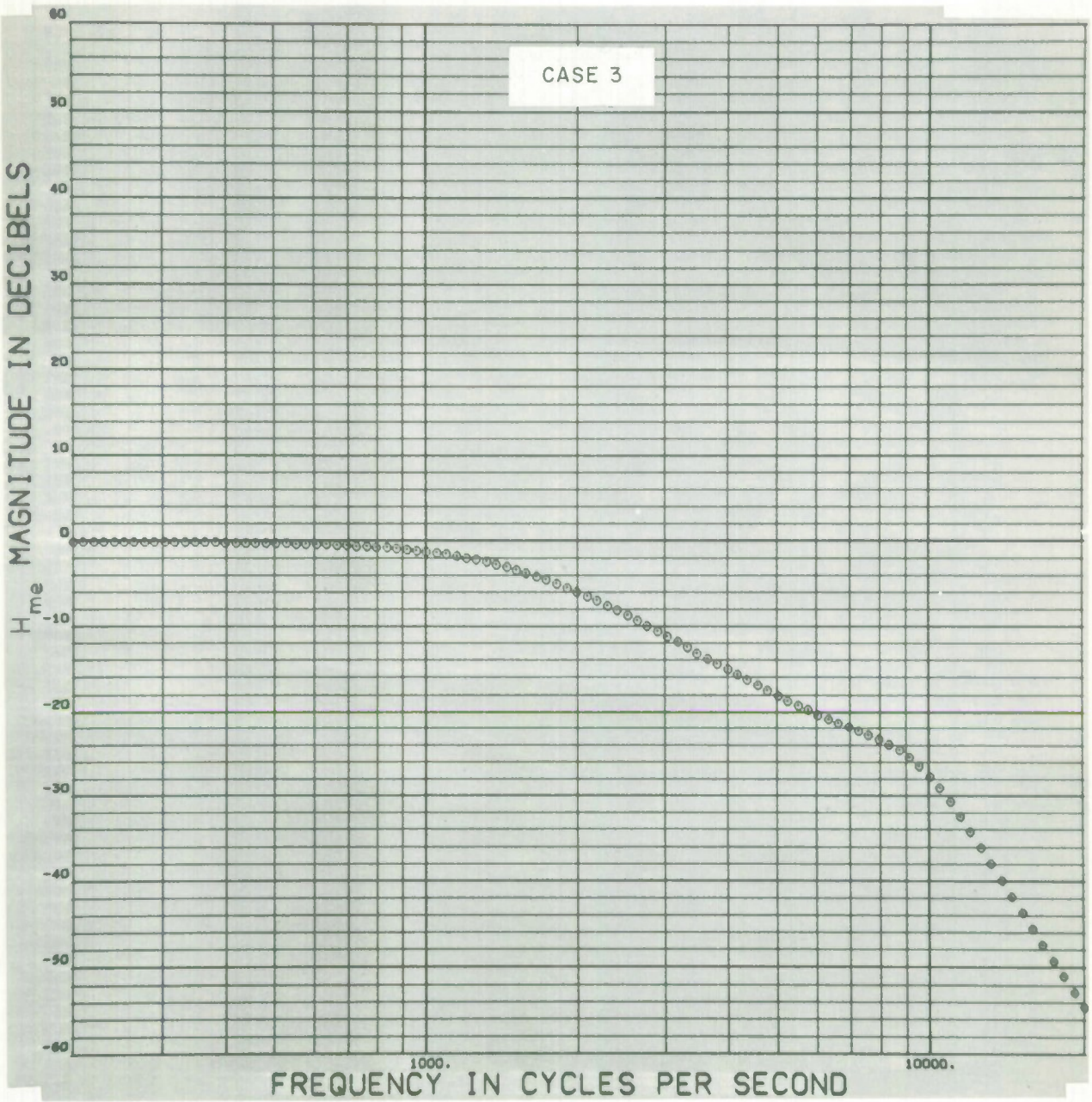

Fig. 3-11. Peake's middle ear circuit analog ($\rho = 1/3, \sigma = 0$).

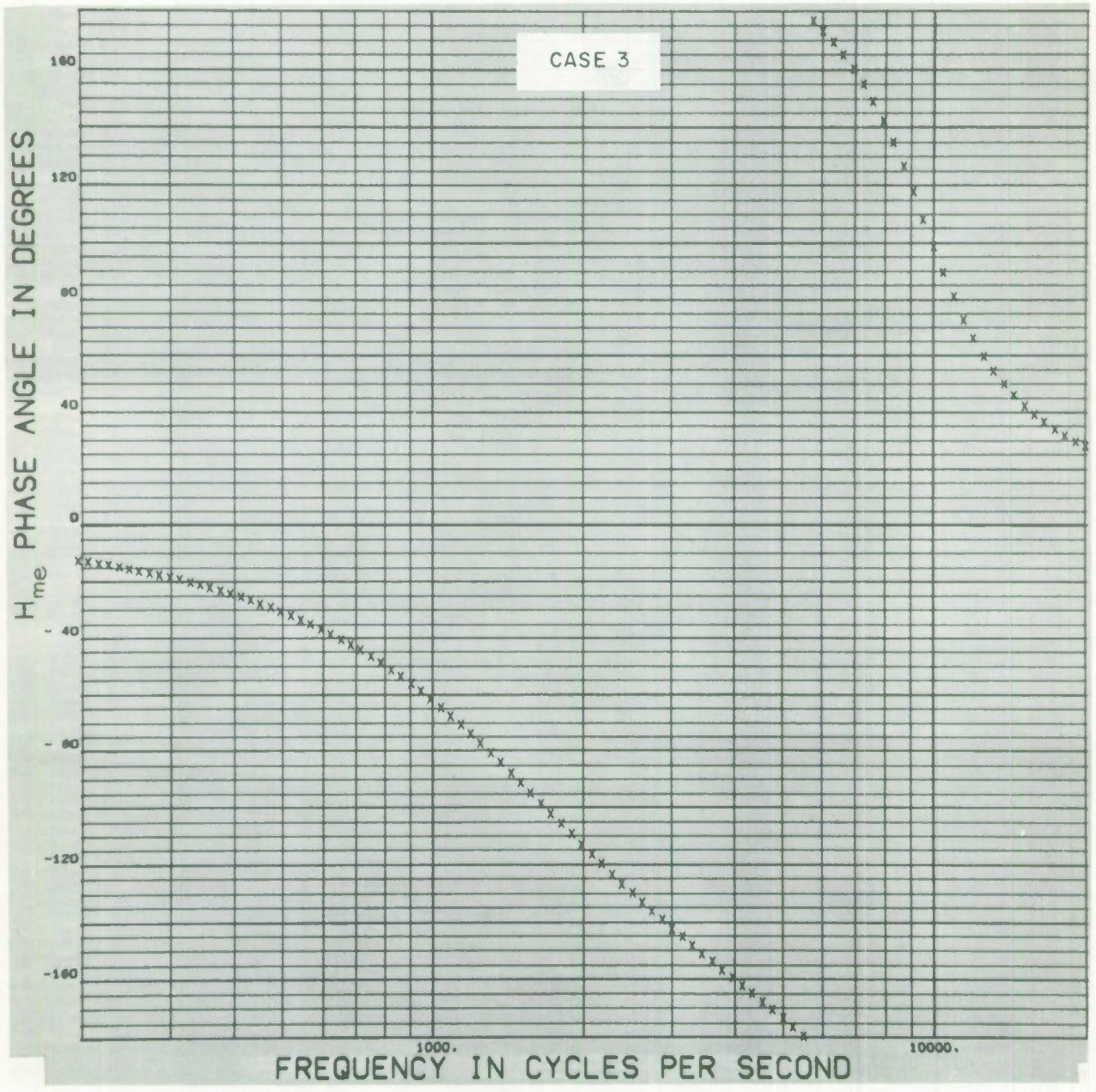

Fig. 3-11. (Continued).

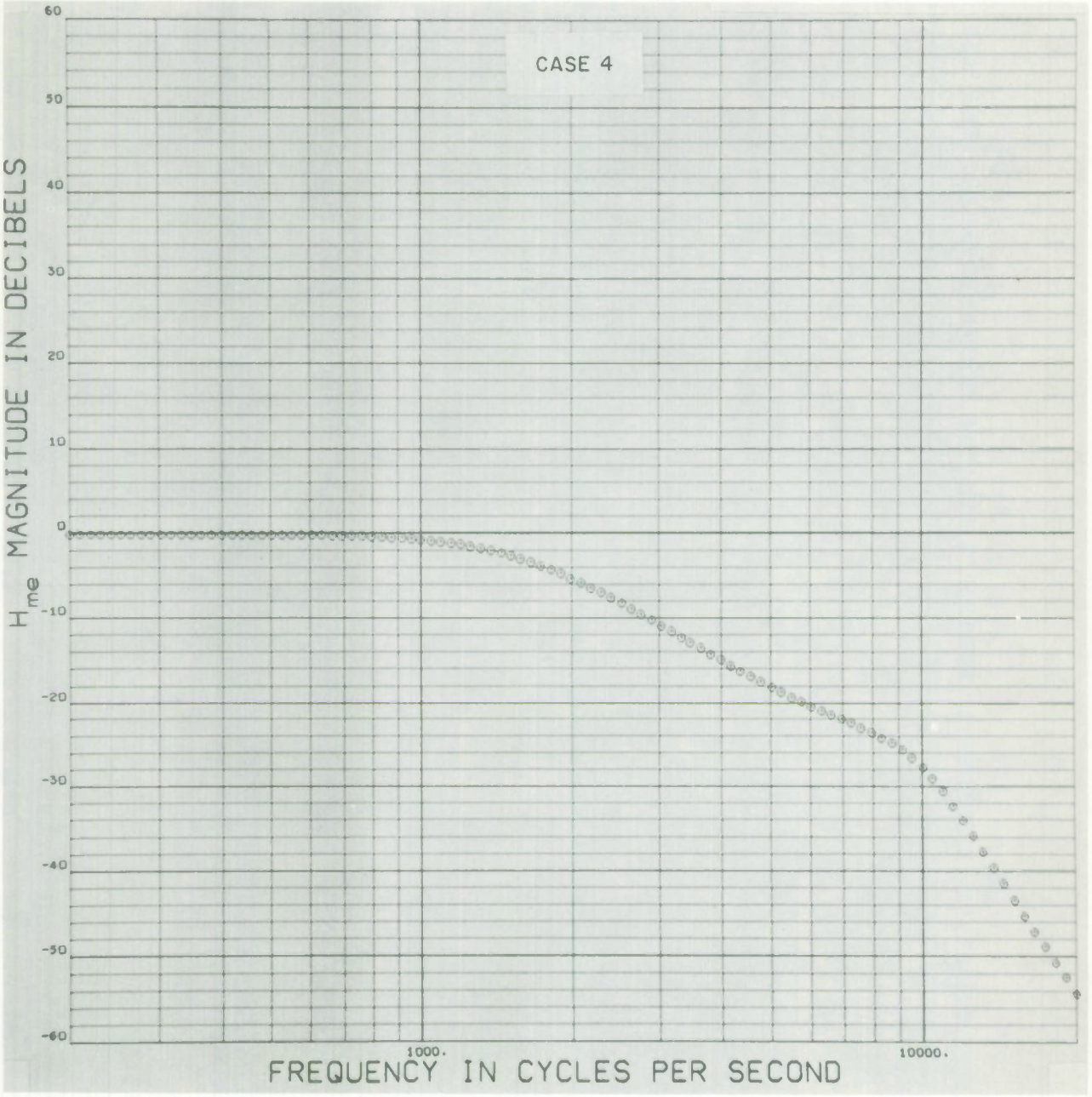

Fig. 3-12. Peake's middle ear circuit analog but with $C_1 = C_2$ ($\rho = 1/3$, $\sigma = 0$).

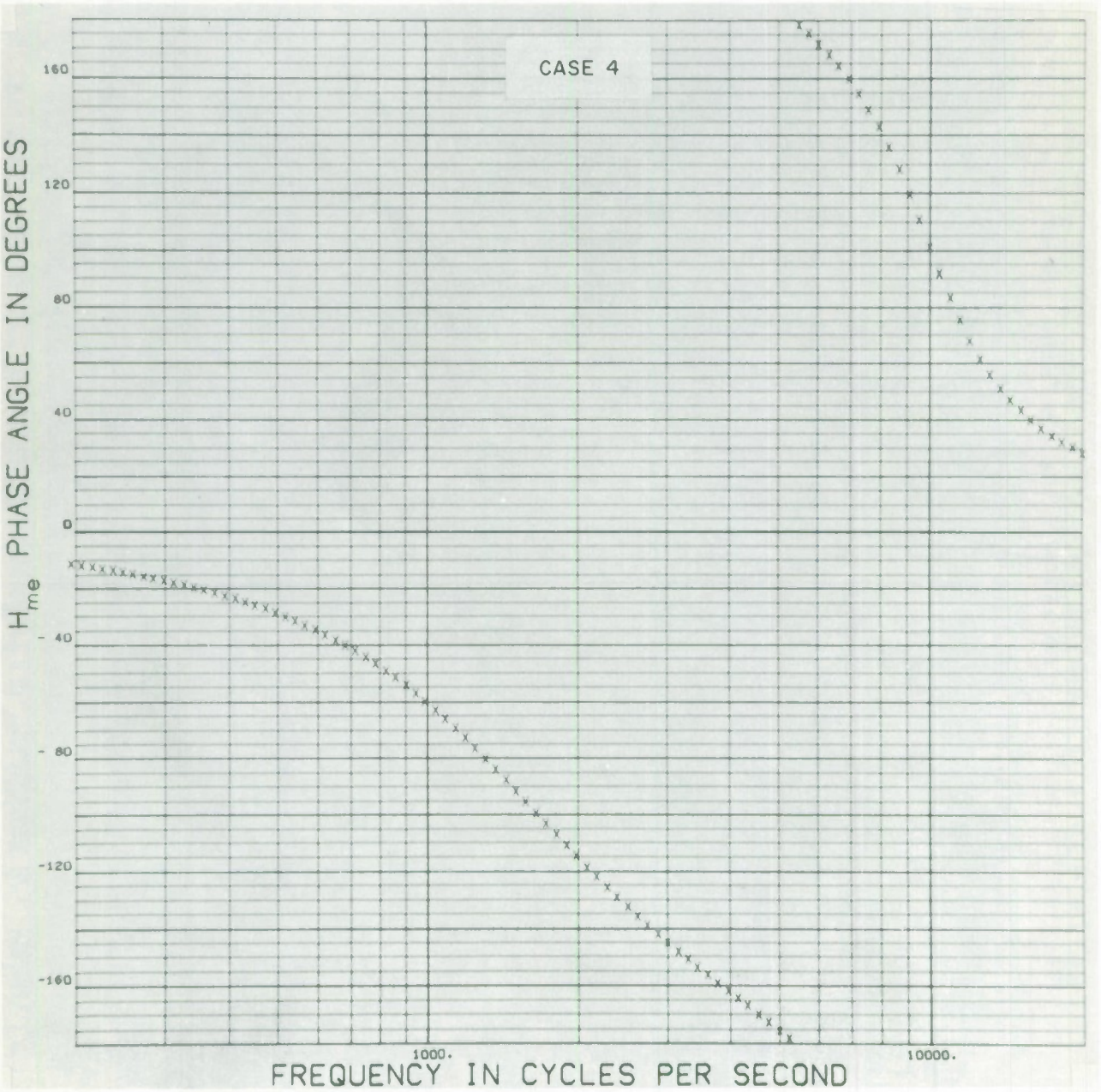

Fig. 3-12. (Continued).

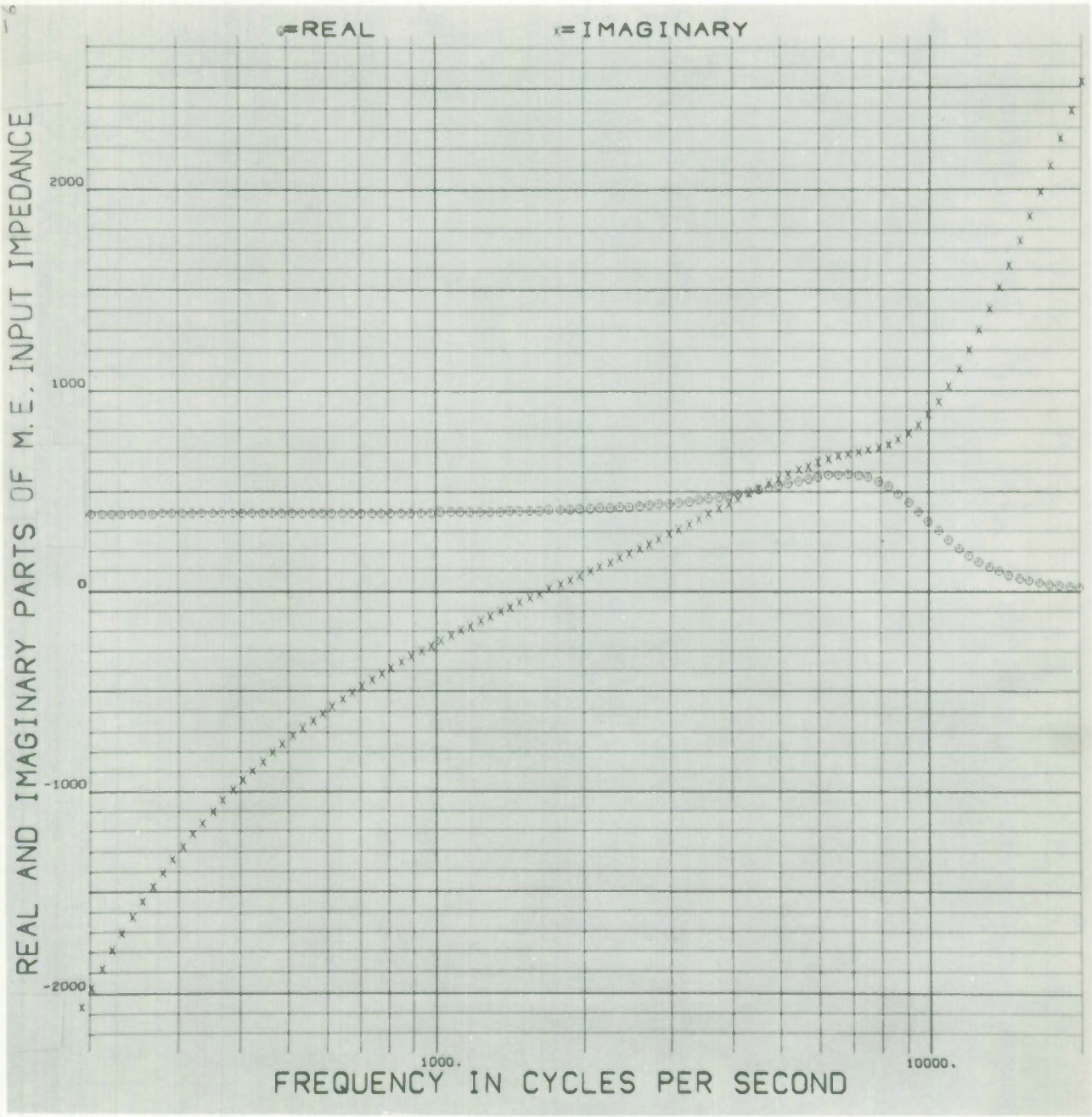

Fig. 3-13. Input impedance of the middle ear circuit analog of Fig. 3-11.

infer a model of the middle ear transfer function from any data other than direct measurements of this transfer function. (Attempts have been made to obtain this transfer function from mechanical properties of the middle ear together with measurements of middle ear input impedance made at the eardrum.)

The middle ear cavities, which are in series with the eardrum and ossicles circuits, have not been included in our model. These cavities are all small compared to a wavelength in air for frequencies up to 4 kHz, so they could be modeled with lumped circuit elements. One would expect the general impedance level of these air-filled cavities to be lower than that of the more massive mechanical system of the ossicles, except perhaps at resonances of the cavities. Unfortunately, there are resonances of the middle ear cavities in humans in the 0- to 4-kHz range, although precise measurements concerning them are scant. Since the middle ear circuit model matches the measured transfer function adequately for our purposes, we will continue to neglect the middle ear cavities.

A comparison of the measured middle ear input impedance (Fig. 3-2),

$$Z_{me}(s) = \frac{P_d(s)}{U_d(s)}$$

(U_d = eardrum velocity amplitude) with the input impedance of the circuit analog of Fig. 3-11 (shown in Fig. 3-13) shows a surprisingly good agreement in view of our neglect of the middle ear cavities. The impulse response of this model is shown in Fig. 3-14.

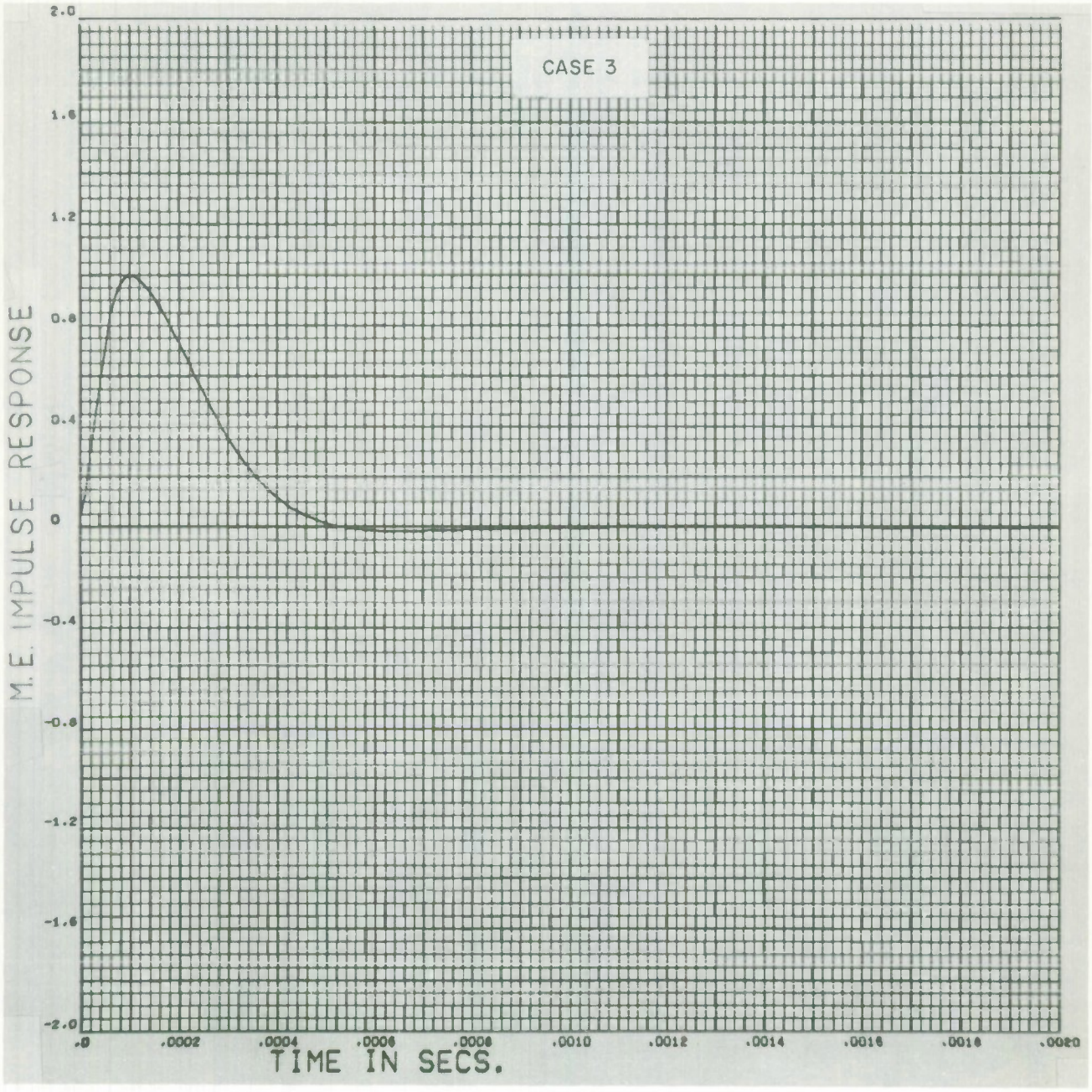

Fig. 3-14. Impulse response (normalized) of the middle ear circuit analog of Fig. 3-11.

Summary

1. We adopt the middle ear circuit analog of Peak and Guinan to match the measured middle ear transfer function.

2. There is a noticeable difference between the middle ear transfer functions measured on human temporal bones and anesthetized cats. Hopefully this is not due to post-mortem changes in the middle ear, but this has not been well established.

3. The form of the electrical circuit derived from a lumped mechanical model of the middle ear is quite adequate to model the measured transfer function.

4. The insensitivity of the model transfer function to changes in certain circuit parameters makes it impossible to draw a close correspondence between circuit element values and mechanical properties, on the basis of presently available measurements.

5. This insensitivity of the circuit model also makes it risky to attempt to infer the middle ear transfer function from other than direct measurements of this quantity.

C. INNER EAR (COCHLEA)

The cochlea is the organ which transforms acoustic vibrations into neural activity in the auditory nerve. We are concerned in this chapter only with cochlear mechanics, i. e., with the mechanical motions of the structures of the cochlea in response to acoustic excitation. The neural activity of the cochlea will be discussed in the next chapter. It should be noted that virtually all contemporary models of cochlear action are based on and compared to the measurements and observations of Nobel Laureate Georg von Bekesy. We are not an exception to this rule.

Experimental Data

We begin with some anatomical measurements made on human cochleas (please refer to the physical description of the cochlea given in Chapter II). The variation of dimensions and physical properties along the cochlea is an immediately conspicuous feature of this organ. The width of the basilar membrane as determined by Wever is shown as a function of distance from the stapes footplate in Fig. 3-15. Shown in Fig. 3-16 is the sum of the cross-sectional areas of the two outer scalae. Static measurements of the elasticity of the basilar membrane at various distances from the stapes were made by von Bekesy by measuring the static pressure difference between the outer scalae necessary to produce a maximum displacement of the cochlear partition of 0.01mm. His results are shown in Fig. 3-17. He points out that he found no significant difference between longitudinal and transverse elasticity at three different points by measuring the force at which a calibrated hair bent when pressed into the membrane. The applicability of these hair-bending experiments to acoustic wave propagation in the cochlea is questionable.

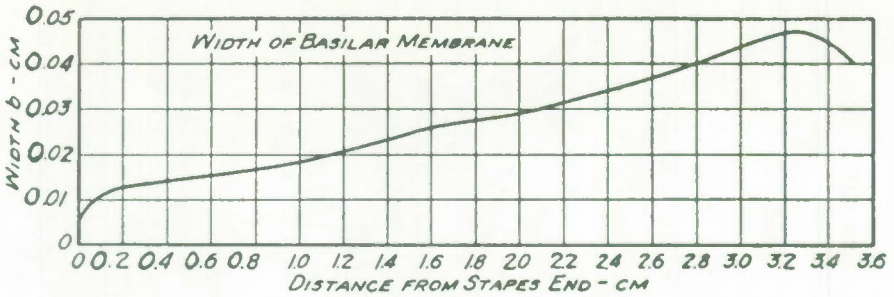

Fig. 3-15. Width of basilar membrane for different distances from the stapes. From Wever, 1938 (in Fletcher).

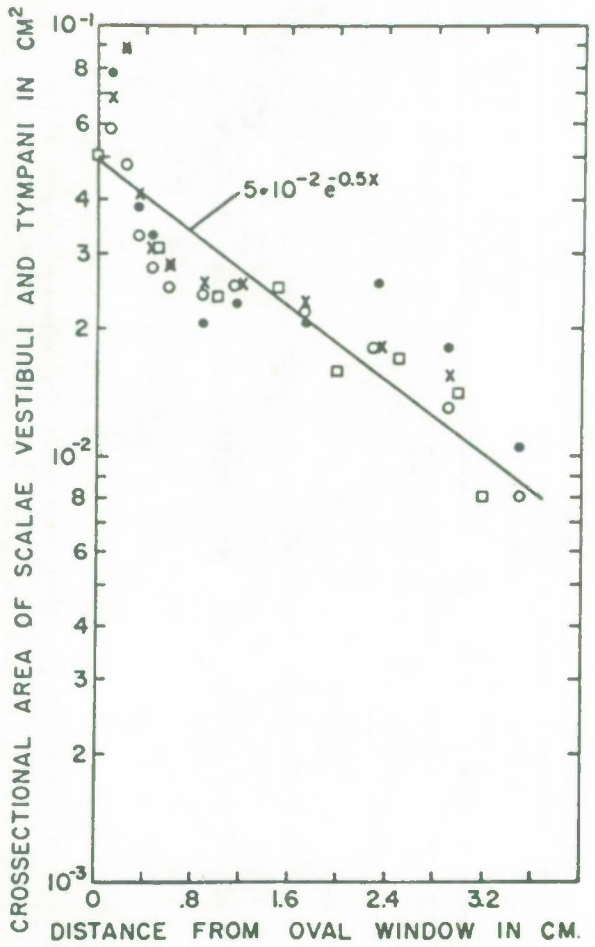

Fig. 3-16. Sum of the cross-sectional area of the scala vestibuli and tympani as a function of distance from the oval window. From Zwislocki (1965), with data from Wever (1949).

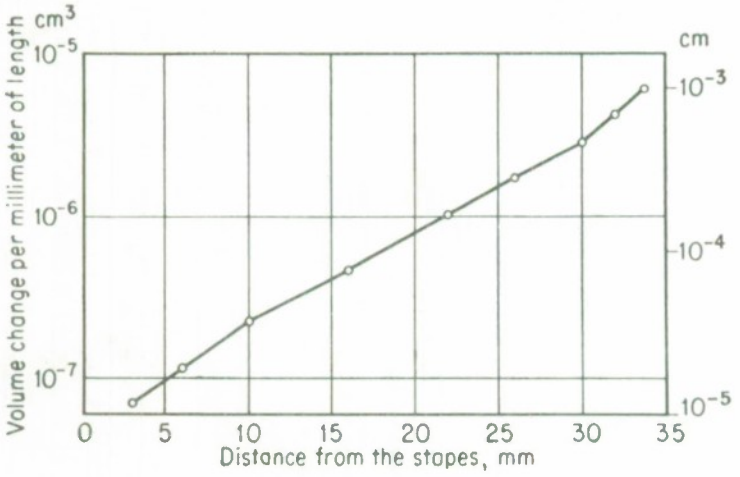

Fig. 3-17. Volume displacement per millimeter of the cochlear duct (left-hand ordinate) and maximum displacement of the cochlear partition (right-hand ordinate) for a pressure of 1 cm of water. From von Bekesy.

The behavior of the cochlea under acoustic stimulation may be described roughly as follows. The footplate of the stapes, projecting through the oval window, acts as a driving piston in the wall of the scala vestibuli. Vibrations of the stapes thus generate an acoustic wave in the fluid of the scala vestibuli. Although the walls of the cochlea are very inelastic and the fluid is almost incompressible, these fluid waves exhibit an interesting mode of propagation. The waves in the scala vestibuli can couple to the scala tympani via the flexible partition between these scalae (consisting of the scala media, the basilar membrane, and Reissner's membrane). The direct fluid path provided by the helicotrema also serves to couple the waves of these two scalae (see Fig. 3-3). The membrane-covered round window acts as a relief valve in the wall of the scala tympani necessitated by the vibrations of the stapes and the incompressibility of the fluids of the cochlea. The organ of Corti, located on the scala media side of the basilar membrane, contains the hair cells which innervate the neurons of the auditory nerve. Vibrations of the basilar membrane induced by the acoustic waves in the scalae fluids then cause a perturbation of the hairs atop these cells by the tectorial membrane which covers them loosely, resulting in neural activity.

The vibration of the basilar membrane for purely sinusoidal sound excitation has been directly observed in human temporal bones. Even for the fairly high sound intensities which were necessary to produce measurable vibrations, the (complex) amplitude of vibration of a point on the basilar membrane was found to be linearly related to stapes amplitude. Denoting the sinusoidal stapes displacement as

$$y_s(t) = \mathrm{Re}\, Y_s(s)\, e^{st}, \quad s = i2\pi f$$

and the displacement of a point on the basilar membrane x cm from the stapes footplate as

$$y_p(t, s) = \mathrm{Re}\, Y_p(s, x)\, e^{st},$$

we define the inner ear transfer function to be

$$H(s, x) = \frac{Y_p(s, x)}{Y_s(s)} . \tag{3-5}$$

We have tacitly assumed that all points on the basilar membrane at a distance of x cm from the stapes footplate vibrate in phase and with a simple scaling of amplitude. This has been observed to be the case; otherwise, $H(s, x)$ would not be sufficient to characterize basilar membrane motions.

In Fig. 3-18 we show the measured frequency response $|H(s, x)|$ vs $f = s/i2\pi$ of several points on the basilar membrane. The curves are normalized so the peak response is unity. It is seen that the frequency sensitivity of the basilar membrane varies with distance from the stapes, i.e., each point along the basilar membrane has a different "best frequency" at which its amplitude of vibration relative to stapes amplitude is greatest. However, the frequency selectivity of points on the basilar membrane is not very great, as can be seen from the width or Q of the frequency response curves.* The best frequencies are plotted in Fig. 3-19 as a function of distance from the stapes [denoted $f_o(x)$], and are seen to decrease with increasing x. The basilar membrane response amplitude at best frequency is shown in Fig. 3-20 plotted

* The Q is defined here as the best frequency divided by the bandwidth at the -3-dB points of $|H(s, x)|$ re its maximum value.

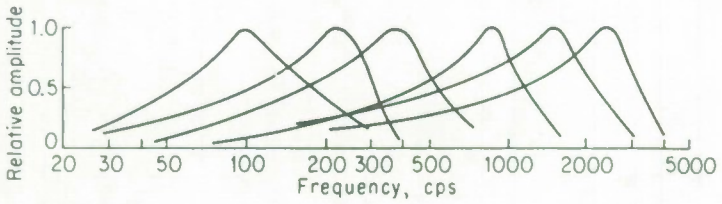

Fig. 3-18(a). Forms of resonance curves for six positions along the cochlear partition. From von Bekesy.

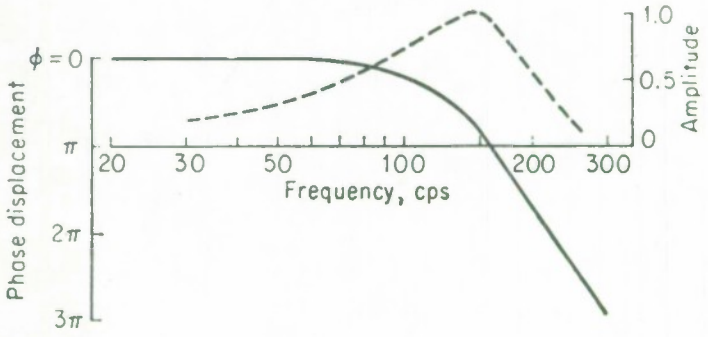

Fig. 3-18(b). Phase displacement (solid line) and resonance curve (broken line) for a point on the cochlear partition 30mm from the stapes. The phase angle is relative to the stapedial motion. From von Bekesy.

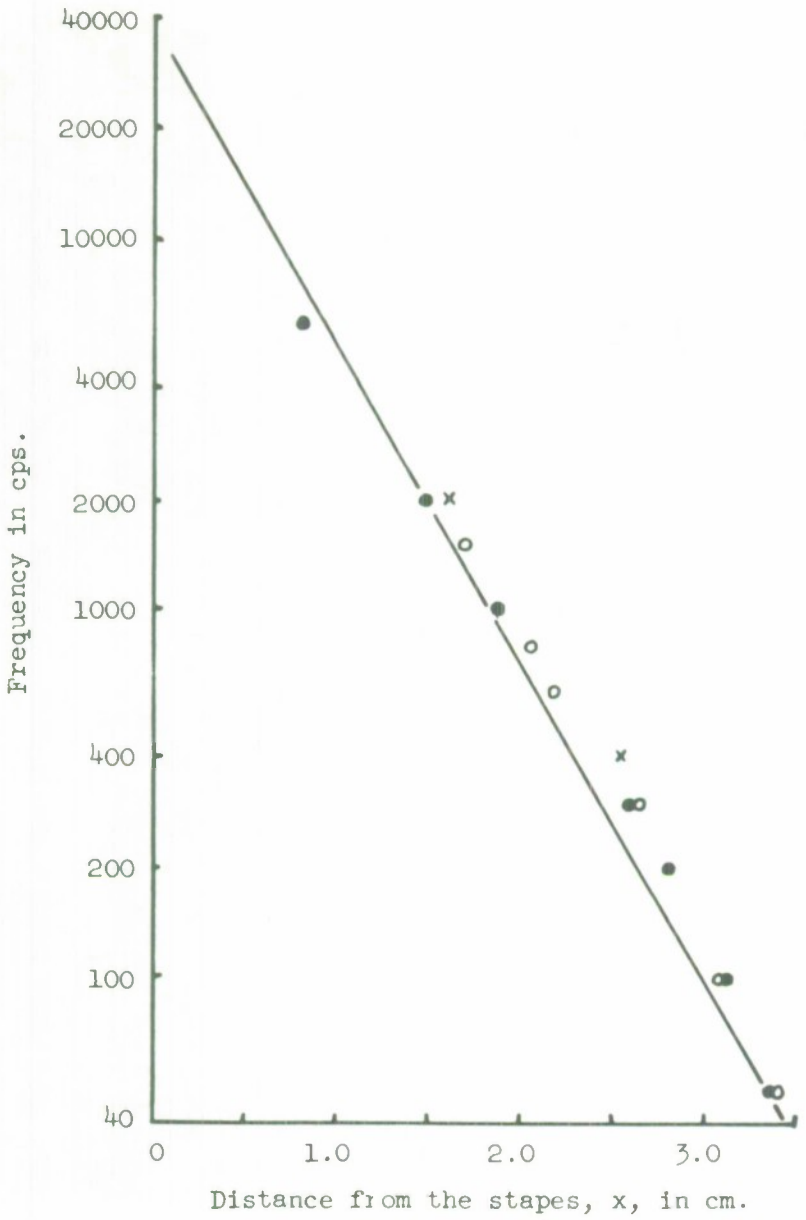

Fig. 3-19. Location of the maximum basilar membrane displacement as a function of the input frequency. The data points are from three experiments of von Bekesy. The straight line is included for reference.

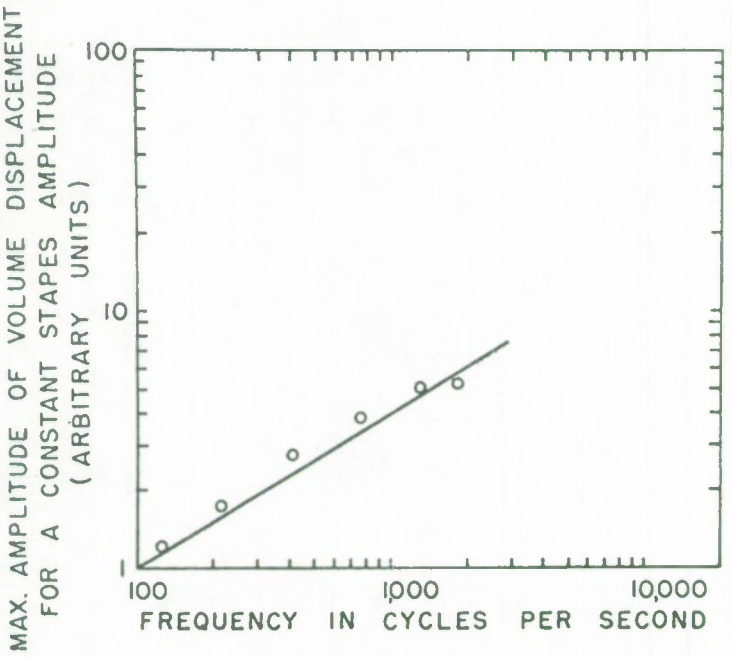

Fig. 3-20. Maximum amplitude of vibration of the cochlear partition for a constant stapes amplitude as a function of sound frequency. From von Bekesy.

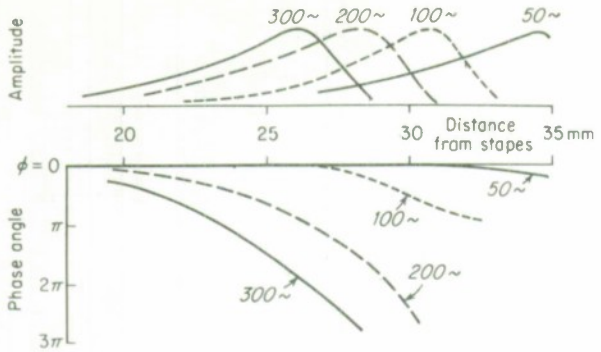

Fig. 3-21(a). Phase displacement and resonance curves for four low tones. From von Bekesy.

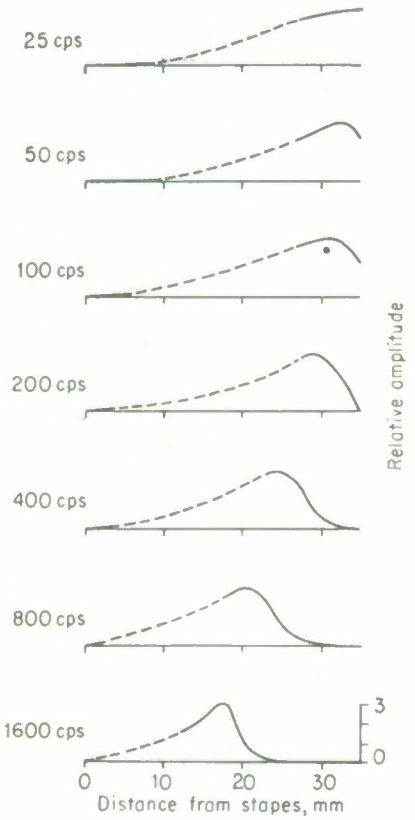

Fig. 3-21(b). Patterns of vibration of the cochlear partition of a cadaver specimen for various frequencies. From von Bekesy.

against best frequency and against distance from the stapes. The maximum responses decrease with distance from stapes. Curves of the magnitude and phase of $H(s, x)$ vs x are shown in Fig. 3-21 for several different frequencies.

One further measurement of interest is the propagation time of a force impulse applied to the stapes to various points along the basilar membrane. However, some arbitrary measure of propagation time must be adopted. Figure 3-22 shows the interval between the first motion of the excitatory impulse and the time at which a point on the basilar membrane is first observed to move. This data is usually interpreted as indicating a changing propagation velocity along the basilar membrane but the dispersion of the propagating waves obscures such a conclusion.

There are several more experimental observations of cochlear action that are well worth mentioning. Direct observations under stroboscopic illumination indicated that the vibratory patterns of the basilar membrane leading to Fig. 3-21 are not standing waves. Rather, sinusoidal excitation was observed to produce a traveling wave on the basilar membrane. These traveling waves attenuate very rapidly and, at high frequencies, the apical end of the basilar membrane remained completely at rest.

An independent verification of the Q of the frequency response of a point on the basilar membrane was obtained from direct measurement of the decay time of the transient when the sinusoidal excitation was suddenly turned off. These measurements indicated a Q of about 1.4 to 1.8 for points all along the basilar membrane. This is consistent with Fig. 3-18.

Direct observations indicate that the basilar membrane and Reissner's membrane are in phase for frequencies below 3 kHz. Therefore, the structure

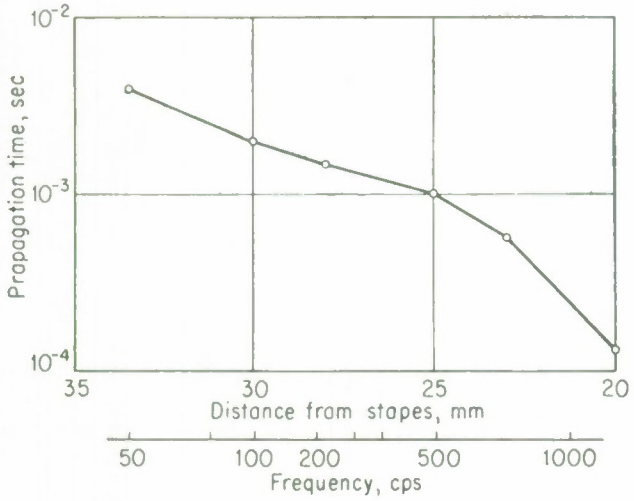

Fig. 3-22. Time of propagation of clicks along the cochlear partition. From von Bekesy.

consisting of the scala media and the basilar and Reissner's membranes acts like a simple flexible partition between the two outer scalae, and so we will refer to this structure as the "cochlear partition."

Some interesting experiments were conducted on models* of the cochlea by von Bekesy in an attempt to determine the sensitivity of the cochlea to various aspects of its structure. He found that the vibratory patterns of the partition in these models were insensitive to increases in the cross sections of the two scalae as well as to the location of the stapes (driving piston) in the wall of the model. However, the partition vibratory patterns did alter significantly when the scalae cross sections were made smaller than actual size (in human cochleas).

Lest one obtain the impression from all of the aforementioned experimental data that cochlear mechanics is a well-understood subject, we will now discuss several troublesome points. Von Bekesy reported observing eddy currents in the scalae fluids around the point of maximum vibration of the partition under sinusoidal stimulation in both his models and in human cochleas. A satisfactory explanation of the significance of these eddies has never been presented. Since the cochlear partition tends to vibrate as a unit, and since the fluid within the scala media is more viscous than that in the other scalae, it seems reasonable to suppose there are no eddies, or at least less prominent ones within the scala media, but this has never been experimentally confirmed.

---

* These were glass-walled structures incorporating two water-filled channels separated by a rubber partition with dimensions and elasticity corresponding to those of the basilar membrane.

In fact, the details of the relative motion between the tectcrial and reticular membranes within the scala media under actual acoustic stimulation are not fully understood. This is perhaps the most conspicuous gap in experimental data concerning cochlear action since these membranes, and in particular, the hair-like projections of the hair cells of the organ of Corti are presumed to be at the site of the mechanical-to-neural transduction.

Observations of the basilar membrane vibratory patterns were all made on the temporal bones of human cadavers. The full extent of post mortem changes in cochlear action are thus not known. Experiments on anesthetized animals would be very valuable in this regard as well as in independently corroborating von Bekesy's observations of cochlear vibratory patterns.

Transducer Mechanics

The vibrations of the cochlear partition serve to drive a transducer which, by electro-chemical action, causes primary neuron discharges which convey information to the brain about the auditory stimulus. While no direct evidence is available concerning the actual site of transduction, it is generally believed to take place in the hair cells of the Organ of Corti. The mechanical motions of the basilar membrane are presumed to be communicated to the hair cells by mechanical perturbation of the hair-like projections atop these cells.

Von Bekesy made observations on a guinea pig of the relative motion between the reticular lamina and the tectorial membrane for sinusoidal acoustic stimulation. Some of the various possible relative motions are indicated in Fig. 3-23 to be dependent upon the curvature of the partition in the longitudinal direction, as well as its displacement from equilibrium.

Davis demonstrates graphically in Fig. 3-24 the possibility of a shearing action between the reticular and tectorial membranes, producing a bending of the hairs atop the hair cells. We present an excerpt from Davis on the fine movements of the organ of Corti.

"The fine movements of the organ of Corti and the tectorial membrane have been observed under the microscope by stroboscopic illumination and described in some detail by Bekesy. In any one segment the basilar membrane, organ of Corti, tectorial membrane, and usually Reissner's membrane also, move in phase with one another (see Fig. 2-4). The basilar membrane is fibrous and elastic, and basically it determines the traveling wave pattern of vibration described above. The cells of Hensen form a soft cushion supporting the stiffer plate of the reticular lamina. The tectorial membrane is hinged like the cover of a book along the edge of the limbus. It is composed of a system of diagonal fibers, and also a jelly-like substance. It yields to slow movements. It returns rather slowly after being detached from its normal attachment to the organ of Corti and then displaced.

"Apparently, as the basilar membrane bulges 'upward' or 'downward' the stiff reticular lamina tends to rock, on the support of the rods of Corti, around an axis at the attachment of the basilar membrane to the bony modiolus. The tectorial membrane swings on its attachment to the limbus. The result is a shearing action between the tectorial membrane and the reticular lamina (see Fig. 3-24). The 'hairs' arise from the cuticular plates of the hair cells, which are set firmly in the reticular lamina, and their outer ends are firmly imbedded in the tectorial membrane; therefore, as the basilar membrane

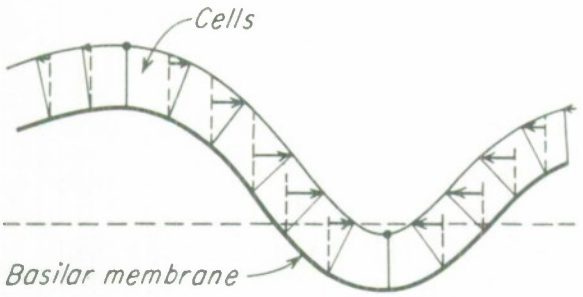

Fig. 3-23. Forms of displacement of the reticular membrane caused by movement of the basilar membrane. From von Bekesy.

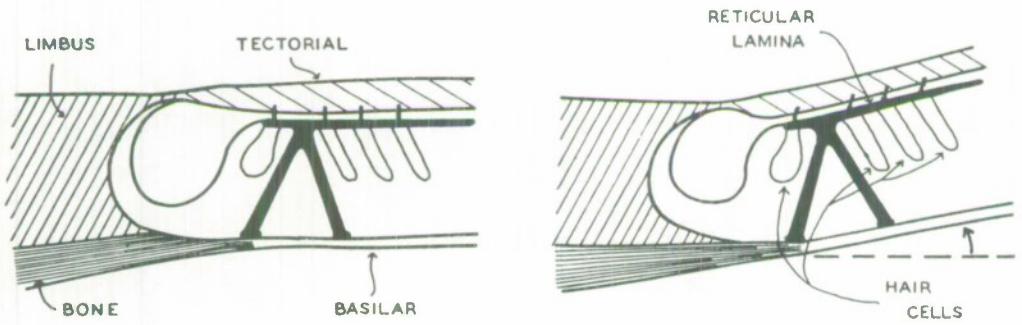

Fig. 3-24. Probable pattern of shearing action between tectorial membrane and reticular lamina, with bending of the hairs. From Davis (1958).

bulges, the hairs are bent. The force of the movements of the cochlear partition is rather efficiently concentrated on the shearing action.

"The movement described above is associated with an approximately radial displacement of Hensen's cells, as seen under the microscope, and a corresponding radial or slightly diagonal bending of the hairs. This movement is characteristic on the basal side of the position of maximal amplitude. On the apical side, however, due to the shorter wave length of the traveling wave and sharper longitudinal bending of the basilar membrane, a longitudinal movement predominates, and the hairs are presumably bent longitudinally instead of radially (see Fig. 3-25).

"The exact significance of these different directions of movement in relation to the excitation of nerve impulses by the hair cells is still a matter of speculation, but the bending of the hairs is the final and critical mechanical event that has been recognized in the mechanism of stimulation. At this point the significant events apparently become electrical, for this bending of the hairs seems to release energy in the form of bio-electric potentials, and these potentials are in all probability an important intermediate step in the mechanism of excitation of the auditory nerve fibers."

Davis (1958)

There are two general points that we wish to make before concluding the discussion of transduction. Firstly, the linearity found elsewhere within the cochlea leads one to expect a linear transformation relating basilar membrane displacement and perturbation (bending perhaps) of the hair projections atop the hair cells, at least for small vibrations. Secondly, the tectorial

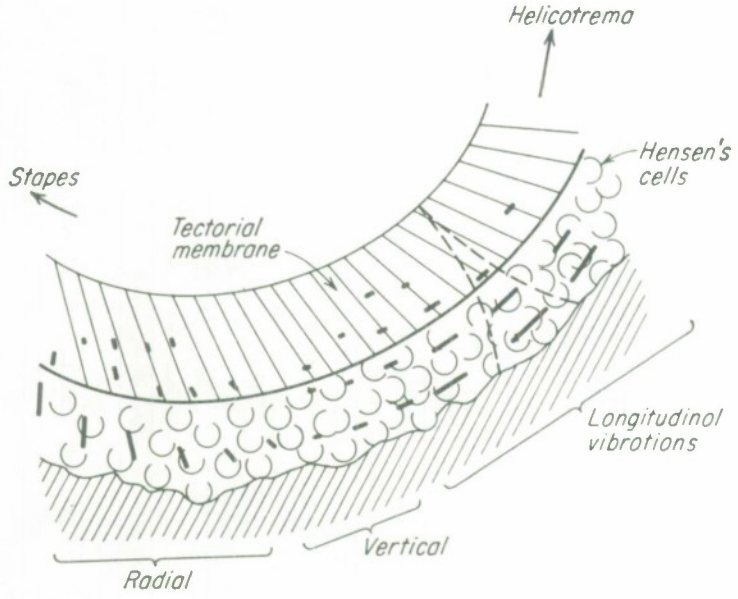

Fig. 3-25. Movements in the cochlear partition, viewed from the upper edge of the stria vascularis. The greatest amplitude of displacement in the pattern shown in Fig. 3-18 is in the region here labeled "vertical." From von Bekesy.

membrane appears to move in phase with the basilar membrane, it is heavily damped by the fluid of the scala media, and there is no evidence of strong elastic forces acting at the hinge of the tectorial membrane. Thus there is no overriding physical reason for expecting an underdamped linear relation between hair cell perturbation and basilar membrane motion.

The evidence indicates that a reasonable possibility for this linear transformation is a memoryless one, i.e., dependent only on the instantaneous values of basilar membrane displacement $y_p(t,x)$, and perhaps its slope, radius of curvature, and higher spacial derivatives. This is pure conjecture, however, and it is not a crucial assumption in what follows.

We thus conclude our discussion of experimental data concerning the action of the cochlea. In the remainder of this chapter we will discuss very briefly several cochlear models, as well as our own first attempts at a model.

Helmholtz' Theory

Helmholtz formulated an early and widely accepted theory of hearing which was based on a frequency analysis of sounds performed by the basilar membrane. He considered the basilar membrane to be composed of taut transverse fibers which were essentially free to vibrate independently. Practically, this effect could be achieved if the longitudinal elasticity of the basilar membrane was much greater than the transverse elasticity. He recognized that,if each fiber could be considered a simple, independent, resonant system, the variation of basilar membrane width, thickness, and transverse elasticity could be sufficient to give rise to transverse fibers with resonant frequencies spanning the entire audible range. Since the cochlear

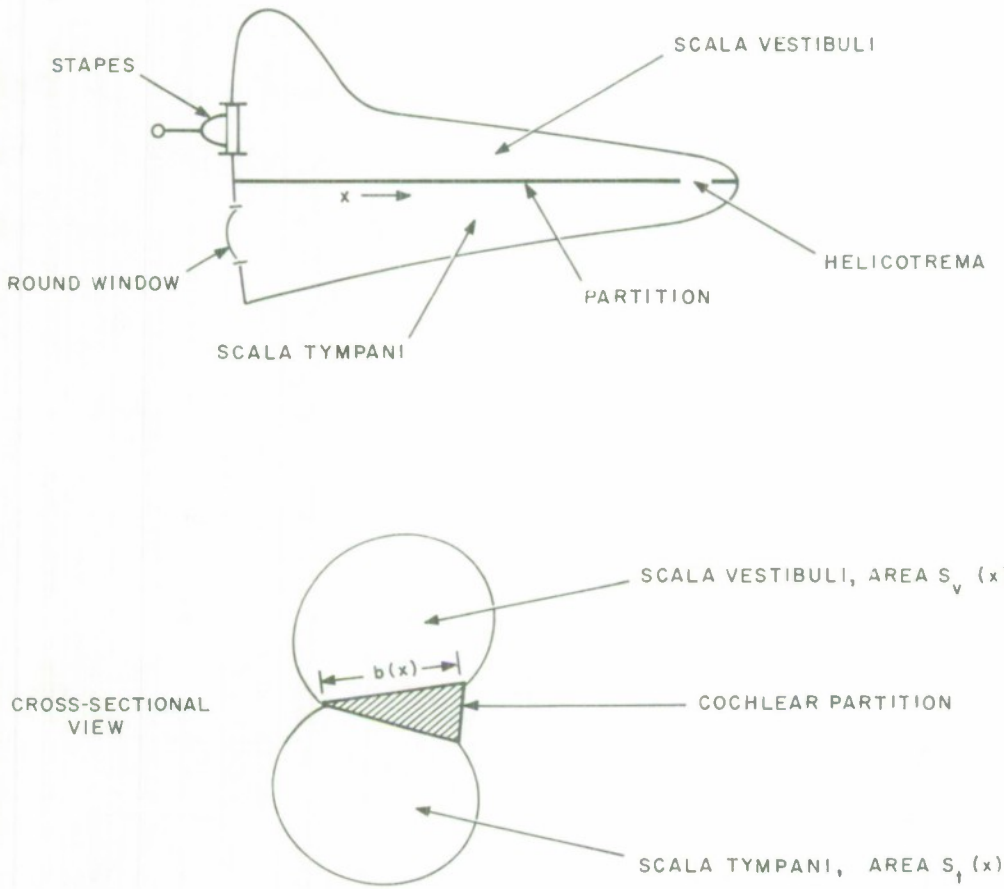

Fig. 3-26. Acoustic model of the uncoiled cochlea.

nerve endings are distributed almost uniformly along the cochlea, he assumed each independent fiber had its own nerve supply.

Helmholtz also recognized that fairly low Q resonators were required in order to be consistent with the then available measurements on discrimination of tones presented in rapid succession. He estimated a Q of about 5 in the region of the basilar membrane which responds maximally at 100 cps.

Of course, von Bekesy's measurements of basilar membrane elasticity do not bear out the anisotropy in elasticity. However, the basic assumptions of the Helmholtz model still permeate many contemporary cochlear models, as we shall soon see.

Contemporary Cochlear Models

It is generally agreed that the coiled arrangement of the cochlea has almost no effect on acoustic wave propagation in the scalae fluid. The uncoiled cochlea is usually modeled as a long, thin, fluid-filled tube divided into two scalae by a flexible partition running the length of the tube (see Fig. 3-26). The cross section of the scalae and the width of the partition vary with distance along the axis of the tube. The cross-sectional dimensions of the scalae are taken to be small with respect to a wavelength in the fluid, resulting in one-dimensional waves in the scalae. Since the perilymph filling the scalae is very similar to water, even the length of the cochlea (3.5 cm) is at least an order of magnitude smaller than the wavelength of a plane acoustic wave in this fluid for frequencies below 4000 Hz.

Assuming sinusoidal excitation, we denote the incremental pressure and average fluid particle velocity in the scala vestibuli (averaged over the scala cross section) as

$$p_v(x,t) = \mathrm{Re}\, P_v(x)\, e^{i\omega t}$$

$$\bar{u}_v(x,t) = \mathrm{Re}\, \bar{U}_v(x)\, e^{i\omega t}$$

where x is the distance along the partition from the stapes. The particle velocity is assumed to be essentially only in the x direction (along the scala) and is averaged over the cross section of the scala. The one-dimensional, small signal equations for acoustic waves in the scala vestibuli are then as follows.

Force equation:

$$-\frac{\partial P_v}{\partial x} = i\omega\rho \bar{U}_v + \sqrt{\frac{2\pi\mu\rho\omega}{S_v}}\,(1+i)\,\bar{U}_v \qquad (3\text{-}6)$$

Continuity equation:

$$-\frac{\partial(S_v \bar{U}_v)}{\partial x} = i\omega \frac{2}{3} b\, Y_p + i\omega \frac{S_v}{\rho c^2} P_v \qquad (3\text{-}7)$$

where we again denote the partition displacement as

$$y_p(t,x) = \mathrm{Re}\, Y_p(x)\, e^{i\omega t}$$

with the convention that positive displacements correspond to bulging of the partition into the scala tympani. We use the notation which is standard throughout the literature, taking

$S_v(x)$ = cross-sectional area of scala vestibuli (cm^2);

$b(x)$ = width of cochlear partition (cm);

ρ = mass density of perilymph (gm/cm^3);

c = velocity of sound in perilymph (cm/sec); and

μ = coefficient of shear viscosity of perilymph (dyne-sec/cm^2).

It is worth noting that we have assumed the basilar membrane has a parabolic shape when displaced from equilibrium* (see Fig. 3-27). Therefore, the fluid volume displaced (per unit length) by the displaced membrane is $\frac{2}{3} bY_p$ where b is the actual width and Y_p is the maximum displacement of the membrane. The $\frac{2}{3}$ factor is only approximate and it is ignored in almost all papers.

Even though the fluid adjacent to the partition will have velocity components other than in the x-direction when the partition vibrates, such velocity components have been calculated to be very small and they are usually ignored.

The second term on the RHS of the force equation is a loss term due to viscosity of the fluid. It has been adapted from results on viscous absorption in uniform, circular pipes of radius $[S_v(x)/\pi]^{\frac{1}{2}}$. Since the scala vestibuli and tympani are approximately circular in cross section, the circular pipe result should provide viscosity losses of the correct order of magnitude for the cochlea. Since viscosity results in a particle velocity distribution which is a function of the radial dimensions of the pipe as well as x, a one-dimensional wave equation is obtained by averaging the velocity over the cross section of the tube. Hence, our equations are in terms of $\overline{U}_v$ which is now only dependent upon x.

The second term on the RHS of the continuity equation represents the compressibility of the fluid. In an incompressible fluid, $c \to \infty$ and this term does not appear.

---

*This is the shape of a thin, uniform membrane for small displacements.

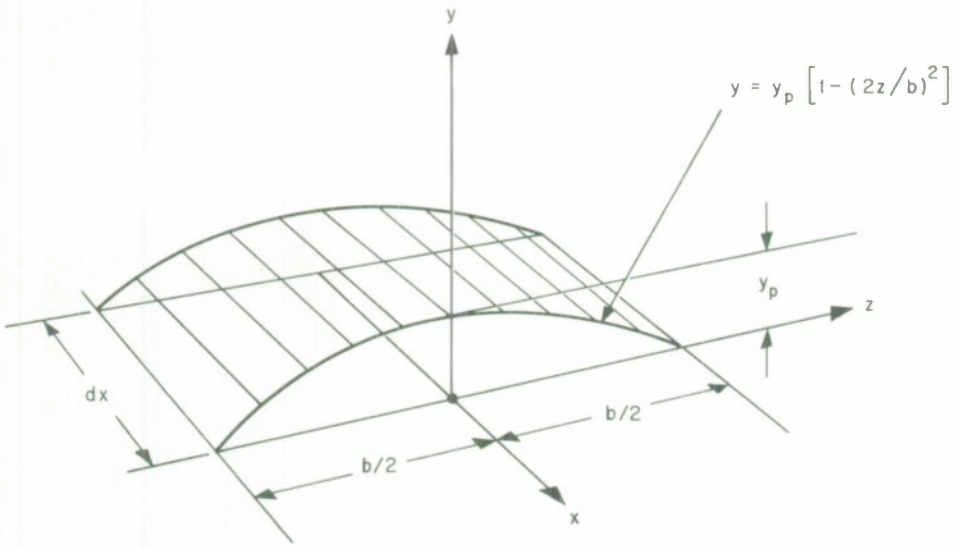

Fig. 3-27. Assumed shape of the cochlear partition when displaced from equilibrium.

There are corresponding equations for the scala tympani which we could write with t subscripts replacing v subscripts (and the sign of the term $i\omega \frac{2}{3} b Y_p$ changed, of course).

The equation of motion of the partition is obtained by considering a segment of the partition dx long and b(x) wide. The force driving such an element of the partition is $(p_v - p_t)b\,dx$ and the differential equation of motion of the partition for $e^{i\omega t}$ excitation becomes

$$(P_v - P_t)b\,dx = K\,dx\,Y_p + i\omega D\,dx\,Y_p + (i\omega)^2 M\,dx\,Y_p + K_L \frac{\partial^2 Y_p}{\partial x^2} \quad (3\text{-}8)$$

where the last term on the RHS represents longitudinal coupling between adjacent segments of the partition. Without this last term, each differential element (along the x axis) of the partition is assumed to have a mass, damping, * and spring stiffness (per unit length) of M(x), D(x), and K(x), respectively, and to behave independently of all other elements of the partition.

There is reason to believe that longitudinal elasticity is a minor effect since the vibratory patterns are so smooth that the partition displacement does not change much in a distance corresponding to the width of the partition. Therefore, the radius of curvature of the partition in the longitudinal direction (proportional to $\partial^2 Y_p/\partial x^2$) is much smaller than in the transverse direction. No analysis yet presented has included this longitudinal elasticity of the partition. Since it does not appear to be an important effect and it does complicate the solution of the equations, we will drop it from our equations.

---

* The partition damping coefficient is sometimes referred to as the viscous coefficient but it does not correspond to the loss term of our force equation which is due to viscosity of the fluid.

The helicotrema is an orifice which provides a direct fluid connection between the outer two scalae. It can be modeled as a lumped mass with damping for frequencies at which the size of the orifice is small with respect to a wavelength, which certainly holds for all cochlear structures. The helicotrema provides a boundary condition linking the waves on the acoustic transmission lines (the scalae).

Making the assumption that the two scala areas are the same simplifies things considerably, and this is not far from the truth. Hence, if we denote

$$S_o = S_v = S_t$$

$$u = \frac{U_v - U_t}{2}$$

$$V = \frac{P_v - P_t}{2} \tag{3-9}$$

the above equations can be written in the following form.

Force equation:

$$\frac{\partial V}{\partial x} = -i\omega\rho u - \sqrt{\frac{2\pi\mu\rho\omega}{S_o}}\,(1+i)u \tag{3-10}$$

Continuity equation:

$$\frac{\partial(S_o u)}{\partial x} = -i\omega\,\frac{2}{3}\,b\,Y_p - i\omega\,\frac{S_o}{\rho c^2}\,V \tag{3-11}$$

Partition equation:

$$V = \frac{1}{2b}\,(i\omega M + D + K/i\omega)\,i\omega\,Y_p \tag{3-12}$$

The parameters of our cochlear model are thus ρ, c, μ, $S_o(x)$, b(x), K(x), M(x), and D(x).

When one neglects longitudinal elasticity in the cochlear partition, each element of the partition is modeled as an independent mechanical system with a simple resonance. In many cochlear models the partition parameters K(x), M(x), and D(x) are chosen so that the resonant frequencies of these elements vary along the partition in agreement with the observed best frequencies as a function of distance given in Fig. 3-19. Although the elements of the partition are damped pistons instead of taut strings, these resonant models bear a very definite resemblance to the Helmholtz theory.

Cochlear models can usually be categorized as "resonance models" or "dispersion models." The resonance models achieve the variation of frequency sensitivity along the cochlea by a resonance of small elements of the partition. The dispersion models do not rely on such a resonance but, instead, on the frequency dependence of the mode of propagation of waves within the cochlea which arises from the dissipative character of the cochlea and its tapering characteristics. In other words, the cochlea may also be modeled as a highly dispersive transmission line to achieve the observed frequency analysis. We will now discuss very briefly some contemporary cochlear models.

Peterson and Bogert

A lossless model of the cochlea was analyzed, i.e., both the coefficient of viscosity μ and the partition damping D(x) were taken to be zero in this model. Both $S_o(x)$ and b(x) were taken to be linear functions of x corresponding to measured dimensions of cochleas. The elasticity K(x) was taken as an exponential function of x corresponding to von Bekesy's determination

of partition elasticity by hair-bending experiments. M(x) was taken to be independent of x and to correspond to the fluid of the scala media. This is an example of a resonance model, and a lossless one at that. The maximum values of the partition amplitude $Y_p(x,\omega)$ turned out to be infinite and so the calculated vibratory patterns of this early model were not realistic. However, reasonable results were obtained from the model in predicting the location of maximum partition displacement as a function of frequency as well as the propagation time of pulses along the partition.

Bogert

The partition damping D(x) was included in this refinement of Peterson and Bogert's earlier work without regard to the physical mechanisms involved. D(x) was taken to be an exponential function of x which presumably resulted in the proper Q of the partition elements. The resulting vibratory patterns of this model were much more realistic.

Fletcher

The equations of Fletcher's model agree exactly with our Eqs. 3-10, -11, -12. He divided the basilar membrane into 140 independent pistons and used the partition elasticity as determined from hair-bending experiments. The partition mass M(x) was taken to be $1.75\ b^3(x)$, which was to have included the effects of the vibrating fluid adjacent to the basilar membrane. The resulting resonant frequencies of the partition elements corresponded closely to the observed frequency localization given by $f_o(x)$. He then chose D(x) simply to correspond to von Bekesy's measured partition attenuation constant. Thus his model appears to be a resonance model. However, he

included losses due to fluid viscosity which could make this a dispersion model. After some small terms were dropped from the equations, solutions were gotten by a numerical integration technique. Unfortunately, the approximations and method of solution gave little insight into the workings of the model. But calculated results agreed well with the measured pulse propagation time along the partition and the vibratory patterns $|H(s,x)|$ vs x. His calculations of the maximum partition amplitude as a function of frequency did not agree with measured results. Fletcher's paper is not easy to follow but his model is a convincing one.

Zwislocki

We refer only to the most recent (to our knowledge) of Zwislocki's (1965) many papers on cochlear mechanics, in which he took the fluid to be incompressible but included losses due to fluid viscosity. The partition stiffness K(x) was taken to correspond to von Bekesy's static measurements. But the choice of M(x) and D(x) definitely did not result in resonant partition elements which in any way approximated the distribution of best frequencies $f_o(x)$. Therefore, this cannot be a resonance model. All model parameters were fitted by functions that varied exponentially with x, allowing an explicit mathematical solution of the wave equations for $e^{i\omega t}$ excitation. The results presented seem to indicate a good match with measured data. The location and amplitude of the maximum partition amplitude as a function of frequency were obtained quite accurately in this dispersion model.

Others

Electronic circuit analogs of the cochlear equations without the loss term due to fluid viscosity were constructed (independently) by Wansdronk and Klatt.

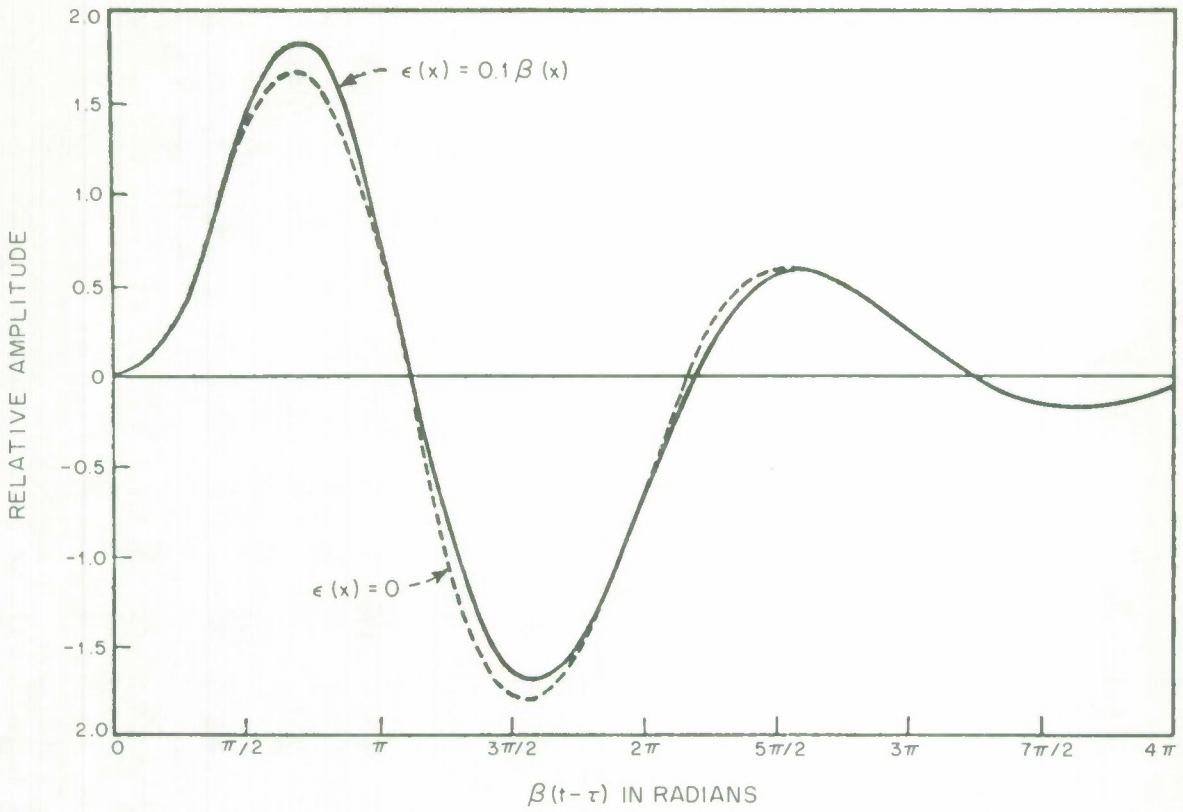

Fig. 3-28. Responses of the basilar membrane model to an impulse of stapes displacement. From Flanagan.

Hause carried out a digital computer simulation of these same equations. These were all resonance models which based the partition elasticity on the hair-bending experiments. The results of these efforts indicate that it is possible to choose plausible values for K(x), D(x), and M(x) which achieve a good fit to the data of Fig. 3-21, even when the fluid is assumed to be incompressible and non-viscous.

Flanagan

The frequency response curves H(s, x) vs s have been fitted by rational functions of s by Flanagan. His approximation to H(s, x) is of the form

$$H(s,x) \approx G\,\frac{s+\epsilon}{s+\beta}\left[\frac{1}{(s+\alpha)^2+\beta^2}\right] e^{-3\pi s/4\beta}$$

where G, ϵ, α, and β vary with x. One advantage of this form of the transfer function is that it allows one to obtain easily the impulse response of a point on the basilar membrane. The general form of this impulse response is invariant with distance along the basilar membrane and is shown in Fig. 3-28.

Electrical Analog of the Cochlea

First define the variables

$$I(x) = S_o(x)\,u(x) \tag{3-13}$$

and

$$I_p(x) = i\omega\,\frac{2}{3}\,b(x)\,Y_p(x) \tag{3-14}$$

Then the cochlear equations may be re-written as follows.

Force equation:

$$\frac{\partial V}{\partial x} = -i\omega L_s I - Z_s I \tag{3-15}$$

Continuity equation:

$$\frac{\partial I}{\partial x} = -I_p - i\omega C_c V \tag{3-16}$$

Partition equation:

$$I_p = \frac{V}{Z_p} \tag{3-17}$$

We have defined

$$L_s(x) = \frac{\rho}{S_o(x)} \tag{3-18}$$

$$Z_s(i\omega, x) = (1+i)\frac{2\pi\rho\mu\omega}{S_o^3(x)}^{\frac{1}{2}} \tag{3-19}$$

$$C_c(x) = \frac{S_o(x)}{\rho c^2} \tag{3-20}$$

$$Z_p(i\omega, x) = \frac{3}{4b^2(x)}\left[i\omega M(x) + D(x) + \frac{K(x)}{i\omega}\right] = i\omega L_p(x) + R_p(x) + \frac{1}{i\omega C_p(x)}. \tag{3-21}$$

Our electrical analog of the cochlea is, of course, nothing more than an association of V(x) and I(x) with voltage and current distributions on a fictitious electrical transmission line which would give rise to the Eqs. 3-15 through 3-21 for $e^{i\omega t}$ excitation. To obtain a computer solution of these equations it is convenient to first approximate them by difference equations. The voltage and current distributions on the line at two points Δx apart satisfy the following conditions when Δx is sufficiently small with respect to a wavelength on the line and any spacial variation of parameters.

$$V(x) \approx V(x+\Delta x) + (i\omega L_s \Delta x + Z_s \Delta x)\, I(x) \tag{3-22a}$$

$$I(x) \approx I(x+\Delta x) + \left(\frac{\Delta x}{Z_p} + i\omega C_c \Delta x\right) V(x+\Delta x) \quad . \tag{3-22b}$$

These equations lead to the lumped electrical circuit approximation to the transmission line in the interval $(x, x+\Delta x)$ that is shown in Fig. 3-29. To complete our electrical analog, we merely terminate our transmission line with a lossy inductor to model the helicotrema. The complete lumped circuit analog of the cochlea is shown in Fig. 3-30.

It is a simple matter to solve this ladder circuit analog of the cochlea for sinusoidal excitation. We first assume a current I_h at the termination of the ladder (say $I_h = 1.0$ amp), which corresponds to the terminal voltage

$$V_h = (R_h + i\omega L_h) I_h \quad ,$$

Then we must have

$$I_N = I_h + V_h Y_N$$

$$V_N = V_h + I_N Z_N \quad ,$$

In general, having V_{n+1} and I_{n+1} at the n-th section of the line, we compute

$$I_n = I_{n+1} + V_{n+1} Y_n$$

$$V_n = V_{n+1} + I_n Z_n \quad ,$$

Thus, by starting at the termination of the ladder with an assumed current, the entire voltage and current distribution along the ladder can be easily computed. The partition displacement $Y_p(x)$ is then obtained from Eqs. 3-14 and 3-17 as

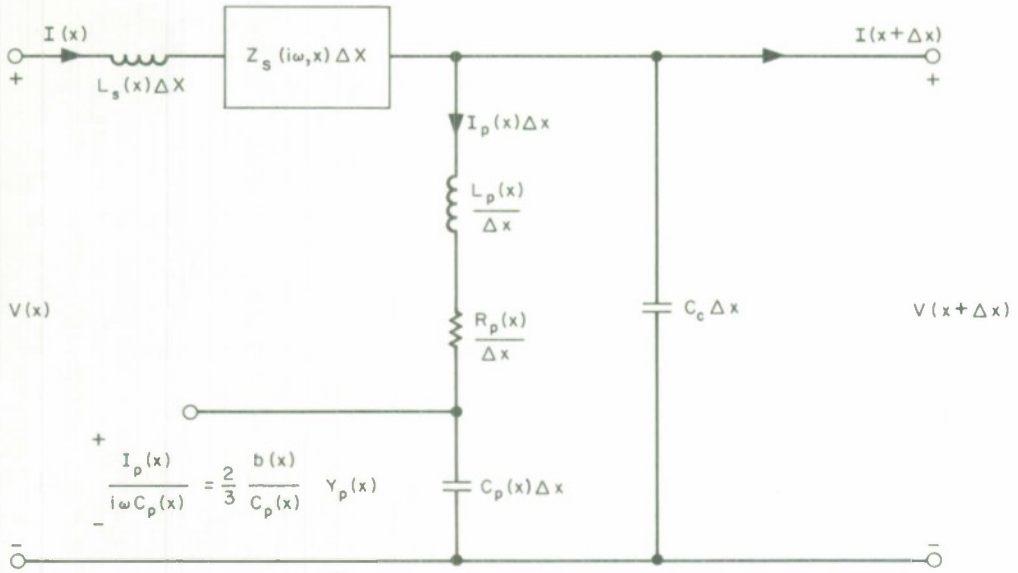

Fig. 3-29. Lumped circuit approximation to segment Δx long of the transmission line analog of the cochlea.

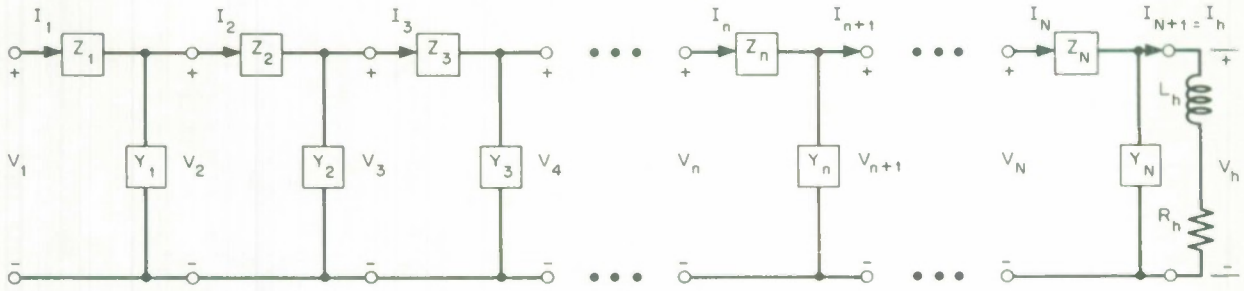

Fig. 3-30. Ladder network analog of cochlea.

$$Y_p(x) = \frac{V(x)}{i\omega \frac{2}{3} b(x) Z_p(i\omega, x)} . \qquad (3\text{-}23)$$

Notice that the voltage across the capacitance $C_p \Delta x$ in the lumped circuit analog (Fig. 3-29) is proportional to Y_p. The transfer function of the cochlea is thus obtained (to within a constant) as

$$H(i\omega, x) = \frac{Y_p(x)}{I(0)} \qquad (3\text{-}24)$$

since $I(0) = I(x = 0)$ will be directly proportional to the stapes displacement.

In Table 3-1 we present our choice of plausible parameter values for a cochlear model. Several choices deserve comment. The parameter $C_p(x)$ can be determined directly from von Bekesy's static measurements of partition elasticity as follows. We can write our partition equation (Eq. 3-8) as it applies to this static test, i.e., $\dot{y}_p = \ddot{y}_p = 0$, as

$$p_v - p_t = \frac{1}{b} K y_p$$

The volume displacement (per unit length) corresponding to the partition motion is $\upsilon = \frac{2}{3} b y_p$ if we assume that the displaced basilar membrane has a parabolic shape (see Fig. 3-27). We can write our static partition equation as

$$p_v - p_t = \frac{3}{2} \frac{K}{b^2} \upsilon = \frac{2}{C_p} \upsilon , \qquad (3\text{-}25)$$

Von Bekesy's static measurements give

$$\upsilon(x)\, dx = 4 \times 10^{-8} e^{1.5x} \text{ cm}^3$$

for $dx = 1$ mm and $p_v - p_t$ equal to a pressure of 1 cm of water (980.7 dynes/cm^2).

Table 3-1 Cochlear Model Parameters

Parameter	Symbol	Units	Values Adopted	Source
Mass density of perilymph	ρ	g/cm^3	1.0	G. von Bekesy
Sound velocity in perilymph	c	cm/sec	1.43×10^5	Peterson and Bogert
Coefficient of shear viscosity of perilymph	μ	$\frac{\text{dyne-sec}}{cm^2}$	2.0×10^{-2}	G. von Bekesy
*Basilar membrane width	b(x)	cm	$e^{-4.42+0.433x}$	Curve fit to data by Wever (1938)
*Total cross sectional area of scala vestibuli and tympani	$S_o(x)^\dagger$	cm^2	$e^{-3.69-0.50x}$	Curve fitted to data by Wever (1949) and Zwislocki (1948) and appearing in Zwislocki (1965)
*Partition compliance (inverse of partition elasticity)	$C_p(x)$	$\frac{cm^4}{\text{dyne}}$	$e^{-20.95+1.50x}$	Curve fitted to static measurement of partition elasticity of G. von Bekesy
*Partition mass per unit length	M(x)	g/cm	neglected	
	$L_p(x)$	g/cm^3	neglected	
*Partition damping per unit length	$R_p(x)$	$\frac{\text{dyne-sec}}{cm^3}$	$\frac{0.39}{2\pi f_0(x)\, C_p(x)}$	
Helicotrema impedance	L_h	g/cm^4	21.2	Morse's equivalent inductance formula for a small acoustic orifice
	R_h	$g/sec\text{-}cm^4$	40,000	

* The parameter is always expressed in cm.

† We make the approximation $S_o(x) \approx \frac{S_v(x) + S_t(x)}{2}$.

Therefore,

$$C_p(x) = \frac{2\times 4\times 10^{-8} e^{1.5x}}{980.7\times 10^{-1}} \approx e^{-20.95+1.5x} \frac{cm^4}{dyne}. \qquad (3\text{-}26)$$

It might seem reasonable to take the mass of the partition to correspond to the fluid of the scala media (as Peterson and Bogert did), since this fluid vibrates almost entirely in the transverse direction. However, this proves to be far too much mass, restricting the partition elements to very low resonant frequencies and giving no frequency localization along the cochlea. From von Bekesy's sketch of the partition deformation (Fig. 3-31), it is clear that all of the fluid particles of the scala media do not move the same amount. In fact, a small fraction of the total fluid of the scala media moves as much as the maximum partition displacement. We, therefore, take the mass of the partition to be negligible in our first cochlear model. Since the partition elements are then non-resonant, this assumption results in a dispersion model of the cochlea.

The partition damping D(x) [or $R_p(x)$] is probably the most difficult parameter to relate to physical reality. If the partition mass is neglected, it can be shown that the maximum of the frequency response of a point on the basilar membrane x cm from the stapes corresponds to the frequency at which the capacitive reactance $1/2\pi f C_p(x)$ of the partition is of the same order as the partition resistance $R_p(x)$. Therefore, we must choose $R_p(x)$ so that

$$R_p(x) \approx \frac{1}{2\pi f_0(x)\, C_p(x)}$$

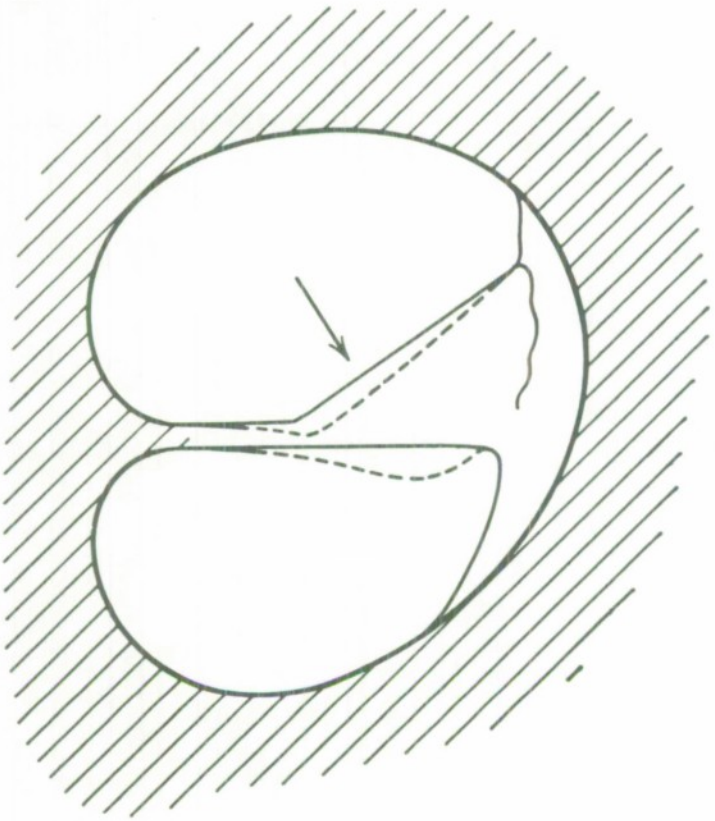

Fig. 3-31. Nature of deformation of the cochlear structures when the stapes was moved inward. From von Bekesy.

if our cochlea model is to exhibit the experimentally determined frequency localization given by the function $f_0(x)$ (see Fig. 3-19). The actual choice of $R_p(x)$ has been slightly refined since the maximum frequency response actually occurs when $R_p(x)$ is somewhere between 0.33 and 0.5 times $1/\omega C_p(x)$, depending on the exact value of x.

The helicotrema can be regarded as a circular orifice in a thin, rigid plate for which the equivalent inductance can be calculated as

$$L_h = 0.8 \frac{\rho}{S_h} \text{ g/cm}^4$$

where S_h is the area of the helicotrema which we take to be 1.4×10^{-3} cm^2 ($\rho = 1$ g/cm^3). The resistive part of the helicotrema impedance was chosen to be equal to the characteristic impedance of the analog transmission line at x = 3.5 cm. This effectively provided a matched termination, reducing wave reflections that resulted in standing waves on the analog line at low frequencies (40 to 140 Hz).

The behavior of our first cochlear model is shown in Figs. 3-32, -33, and -34 and Table 3-2. Comparison of these data with Figs. 3-18, -19, and -20 reveals several shortcomings in our model.

1. The width or Q of the frequency response curves of Fig. 3-32 is larger than that of the measured curves (Fig. 3-18).

2. The pulse propagation times are difficult to calculate for highly dispersive transmission lines. We have calculated the pulse propagation times for various points along the partition neglecting all losses in our model. These propagation times are all longer than the measured ones of Fig. 3-22. Losses would tend to further increase the propagation times.

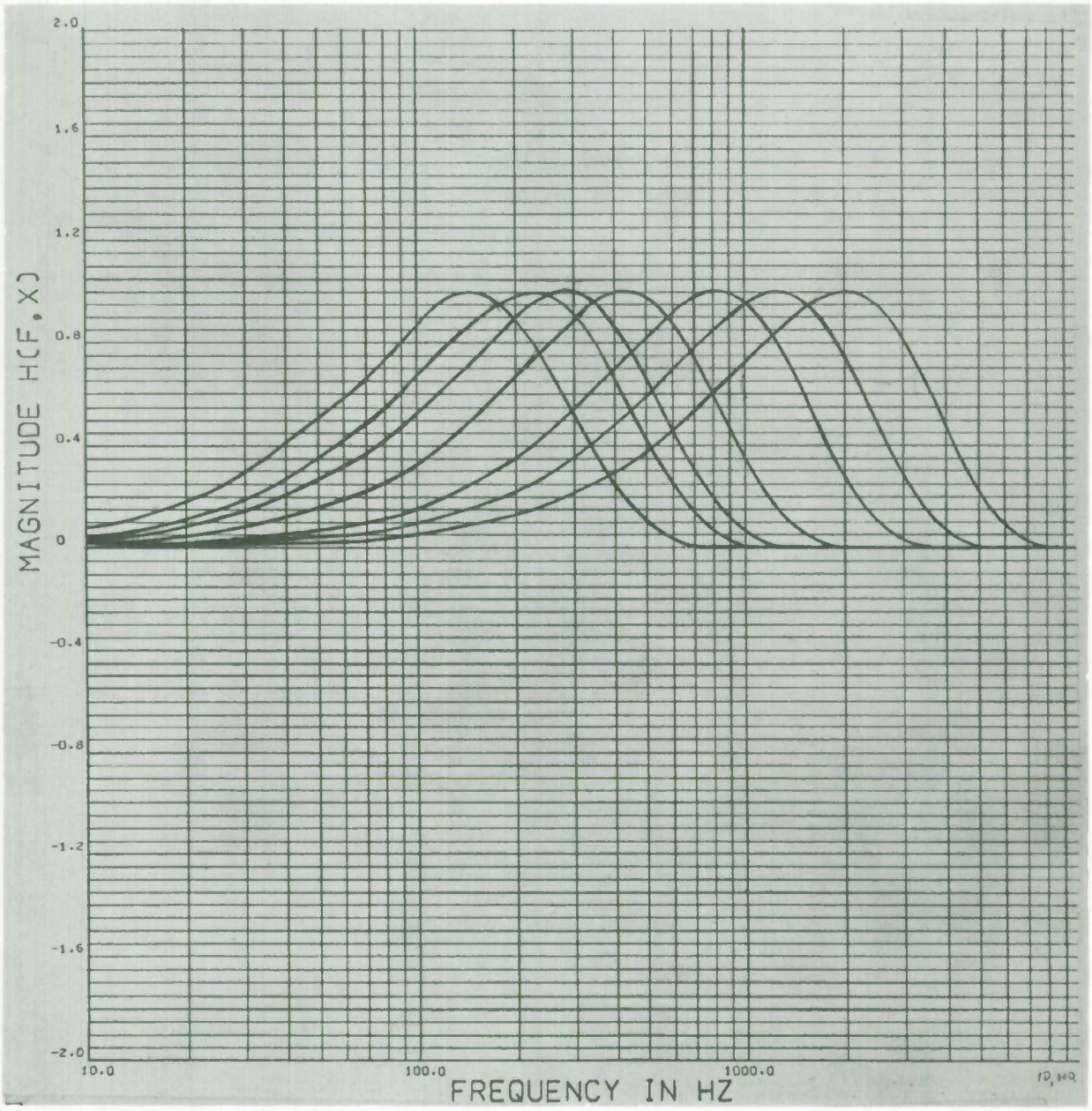

Fig. 3-32. Frequency response of first cochlear model.

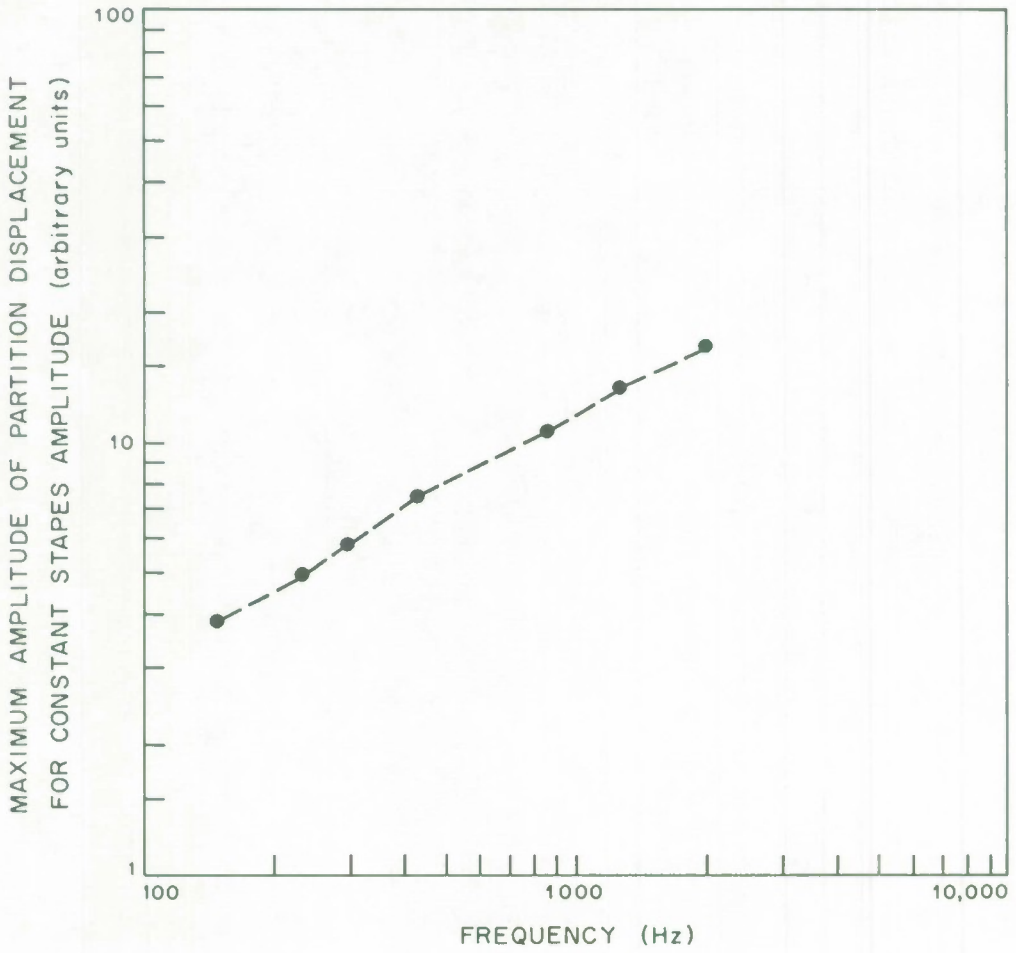

Fig. 3-33. Maximum amplitude of partition displacement of first cochlear model for constant stapes amplitude.

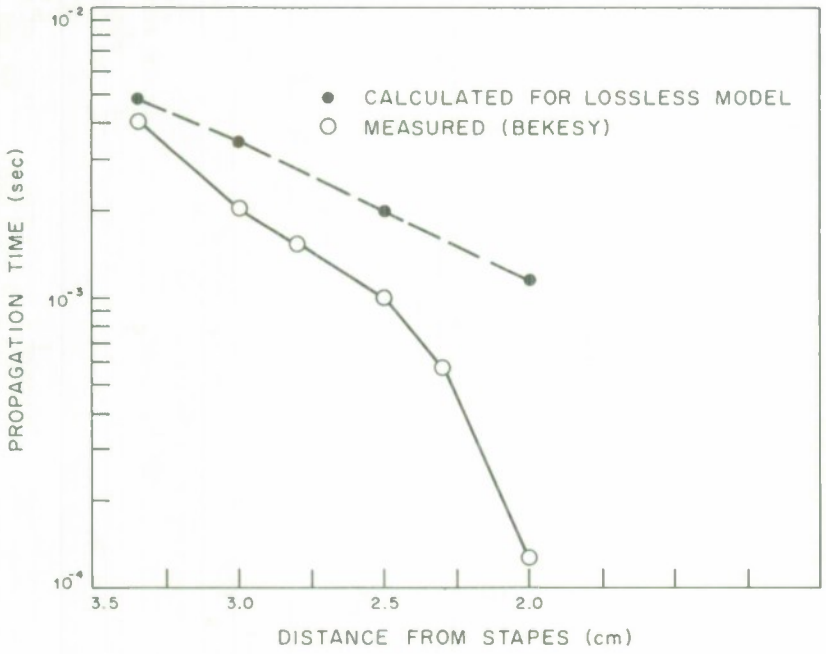

Fig. 3-34. Propagation time for clicks traveling along the cochlear partition for first cochlear model.

Table 3-2

Data from First Cochlear Model

x(cm)	f_{max}	$\|H\|$ rel. stapes (arbitrary units)
1.35	1985	16.8
1.70	1253	13.3
2.00	854	10.7
2.50	428	7.49
2.80	292	5.80
3.00	232	4.95
3.30	146	3.88

3. The high loss attributed to the helicotrema was necessary to match out wave reflections at low frequencies. This high loss is probably physically unrealistic. Therefore, the cochlear model is not lossy enough at low frequencies.

Summary

1. While there is an abundance of data on cochlear mechanics, the data are incomplete in several respects. Insufficient phase angle data in measurements of $H(s, x)$ preclude a clear decision on whether a resonance or a dispersion model of the cochlea is more physical. The effective partition mass and damping are almost totally unknown. These omissions allow some freedom in choosing cochlear models to match the existing data.

2. No cochlear model known to us matches all of the existing data on cochlear mechanics. The general form of $H(s, x)$ is relatively easy to match, and the impulse response $h(t, x)$ seems insensitive to the details of the model.

3. Calculation of the basilar membrane motion for arbitrary acoustic stimuli is a time-consuming task even for a large digital computer. Usually lumped circuit approximations to electrical transmission line analogs are built.

4. The general nature of the perturbation of the hair projections atop the hair cells is understood. Such details as the difference between perturbations experienced by the inner and outer row of hair cells are unknown. The mechanical stimulus that actually drives the transducer is thus unknown.

5. Further experimental work on cochlear mechanics on anesthetized animals would serve to corroborate the available data, enlarge the general fund of knowledge of this subject, and eliminate the effects of post mortem changes in cochlear characteristics.

External Ear References

F. M. Weiner, "On the Diffraction of a Progressive Sound Wave by the Human Head," J. Acoust. Soc. Am. 19, 143-146 (1947).

F. M. Weiner, D. A. Ross, "The Pressure Distribution in the Auditory Canal in a Progressive Sound Field," J. Acoust. Soc. Am. 18, 401-408 (1946).

J. Zwislocki, "Analysis of Some Auditory Characteristics," Handbook of Mathematical Psychology, Vol. III, edited by R. Duncan Luce, R. R. Bush, and E. Galanter (J. Wiley and Sons, 1965).

Middle Ear References

J. L. Flanagan, Speech Analysis Synthesis and Perception (Academic Press, Inc., New York, 1965).

J. J. Guinan, Jr., "The Transfer Characteristics of the Cat's Middle Ear," M. S. Thesis, Department of Electrical Engineering, M. I. T. (1964).

J. J. Guinan, Jr., and W. T. Peake, "Motion of Middle-Ear Bones," Research Laboratory of Electronics, M. I. T., QPR No. 74, 219 (15 July 1964).

A. R. Møller, "Network Model of the Middle Ear," J. Acoust. Soc. Am. 33, 168-176 (1961).

A. R. Møller, "Transfer Function of the Middle Ear," J. Acoust. Soc. Am. 35, 1526-1534 (1963).

A. R. Møller, "An Experimental Study of the Acoustic Impedance of the Middle Ear and its Transmission Properties," Acta Oto-Laryng. 60, 129-149 (1961).

Y. Onchi, "Mechanism of the Middle Ear," J. Acoust. Soc. Am. 33, 794-805 (1961).

W. T. Peake and J. J. Guinan, Jr., "Circuit Model for Transfer Characteristics of the Cat's Middle Ear" (presented at 71st meeting, Acoust. Soc. Am., June 1966), J. Acoust. Soc. Am. 39, 1252, Abs. 6A6 (June 1966).

Middle Ear References (Concl.)

S. S. Stevens and H. Davis, Hearing (John Wiley and Sons, New York, 1938).

J. Zwislocki, "Some Impedance Measurements on Normal and Pathological Ears," J. Acoust. Soc. Am 29, 1312-1317 (1957).

J. Zwislocki, "Analysis of the Middle-Ear Function, Part I: Input Impedance," J. Acoust. Soc. Am 34, 1514-1523 (1962).

J. Zwislocki, "Analysis of the Middle-Ear Function, Part II: Guinea Pig Ear," J. Acoust. Soc. Am. 35,1034-1040 (1963).

Cochlear Mechanics References

B. P. Bogert, "Determination of the Effects of Dissipation in the Cochlear Partition by Means of a Network Representing the Basilar Membrane," J. Acoust. Soc. Am. 23, 151-154 (March 1951).

J. L. Flanagan, "Computational Model for Basilar Membrane Displacement," J. Acoust. Soc. Am. 34, 1370-1376 (1962).

J. L. Flanagan, Speech Analysis Synthesis and Perception (Academic Press, Inc., New York, 1965).

H. Fletcher, "On the Dynamics of the Cochlea," J. Acoust. Soc. Am. 23, 637-645 (1951).

A. D. Hause, Digital Simulation of the Cochlea, Internal Note, Bell Telephone Laboratories, New Jersey (1963).

L. E. Kinsler and A. R. Frey, Fundamentals of Acoustics, First Edition (John Wiley and Sons, New York, 1950).

D. Klatt, "Theories of Aural Physiology," University of Michigan Communications Science Laboratory Report No. 13 (November 1964).

P. M. Morse, Vibration and Sound (McGraw-Hill, New York, 1948).

L. C. Peterson and B. P. Bogert, "A Dynamical Theory of the Cochlea," J. Acoust. Soc. Am. 22, 369-381 (1950).

V. Salman, "Generalized Plane Wave Horn Theory," J. Acoust. Soc. Am. 17, 199-211 (1946).

W. M. Siebert, "Models for the Dynamic Behavior of the Cochlear Partition," RLE QPR No. 64, pp. 242-258 (Massachusetts Institute of Technology, 15 January 1962).

G. von Bekesy, Experiments in Hearing, edited by E. G. Wever (McGraw-Hill, New York, 1960).

C. Wansdronk, "On the Mechanism of Hearing," Philips Research Reports, Supplement 1 (1962).

E. G. Wever, "The Width of the Basilar Membrane in Man," Ann. Otol. Rhinol. and Laryngol. 47, 37-48 (1938).

J. Zwislocki, "Analysis of Some Auditory Characteristics," Handbook of Mathematical Psychology, Vol. III, edited by R. Duncan Luce et al. (John Wiley and Sons, 1965).

CHAPTER IV

The Neural Portion of the Peripheral Auditory System

The final mechanical event in the mechanism of stimulation is the perturbation of the hair projections of the hair cells. Beyond this stage, the significant events are electro-chemical. Unfortunately, the electro-chemical events within the organ of Corti cannot be conveniently monitored during acoustic stimulation. We are required by the availability of experimental data to next consider the neural outputs of the primary auditory neurons. However, before we go into some of the specific properties of the outputs of primary auditory neurons under various stimulating conditions, a very brief description of some of the salient features and properties of neurons in general will be discussed. It is necessary to be acquainted with these inasmuch as many of these features have been incorporated into models of primary auditory neurons that will be discussed below.

Very crudely, a neuron can be considered as a constant-amplitude pulse generator. It is the temporal structure of the pulse train generated by the neuron that carries information about the signal (stimulus) applied to the receptors associated with the neuron. These "information carrying" pulse trains are propagated along the output fiber or axon of the neuron for further processing by neurons elsewhere in the system (e. g. , the output of each primary auditory neuron propagates to an undetermined number of neurons--perhaps hundreds--in the cochlear nucleus).

While one cannot consider all neurons to function in exactly the same manner, a relatively general picture can be representative. The major parts

of a neuron are the cell body (soma), projections from the body (dendrites), and the output fiber (axon). These are enclosed by a membranic wall which is polarized such that a potential difference of 70 to 90 millivolts exists across it (negative inside) when the cell is at rest. The degree of polarization can be altered by noise or as a result of external influences (e. g., receptor action or inputs from other neurons). If the membrane depolarizes sufficiently, the neuron will initiate a regenerative process which results in further depolarization--changing the polarity and amplitude of the potential across the membrane to approximately -20 millivolts outside before reversing the process and returning to the value of the resting potential. The fast portion of this regenerative process or spike discharge is on the order of 1 msec in duration. The spike discharge is essentially fixed in amplitude, i. e., it is either "all or nothing" in character, and the amplitude is independent of the amplitude of the stimulus with other factors held constant. It is the pulse output signal that is propagated along the axon of the neuron (by a traveling region of depolarization) to its (possibly multiple) termination(s) [see Fig. 4-1a). Therefore, we need not be concerned with the waveshapes of the pulses, but only with the times at which pulses are generated.

Several inherent properties of a neuron put a restriction on the temporal structure of the pulse trains that a neuron can generate. The most obvious constraint is that there is a limit on the maximum pulse-repetition rate that can be maintained as a result of a property called refractoriness. Refractoriness is a time-varying state of excitability of the membrane (i. e., a varying depolarization potential necessary to initiate a regenerative discharge)

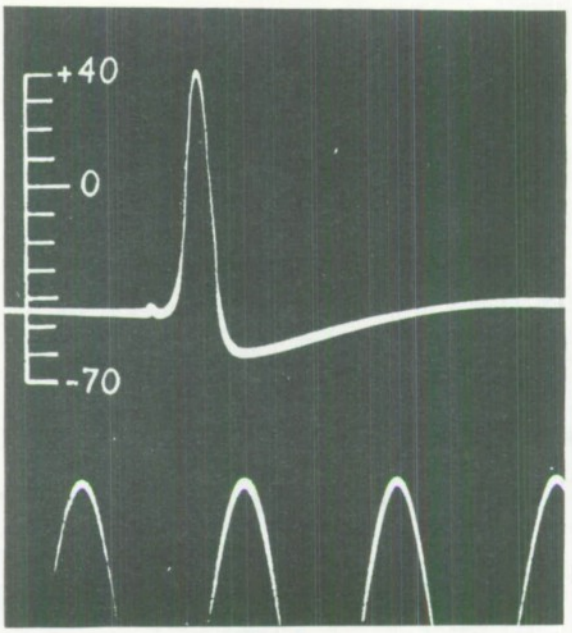

Fig. 4-1(a). Spike discharge potential recorded between inside and outside of axon (squid nerve) in millivolts. The vertical scale indicates the potential of the internal electrode in millivolts, the sea water outside being taken as at zero potential. Time marker 500 cps. In Brazier.

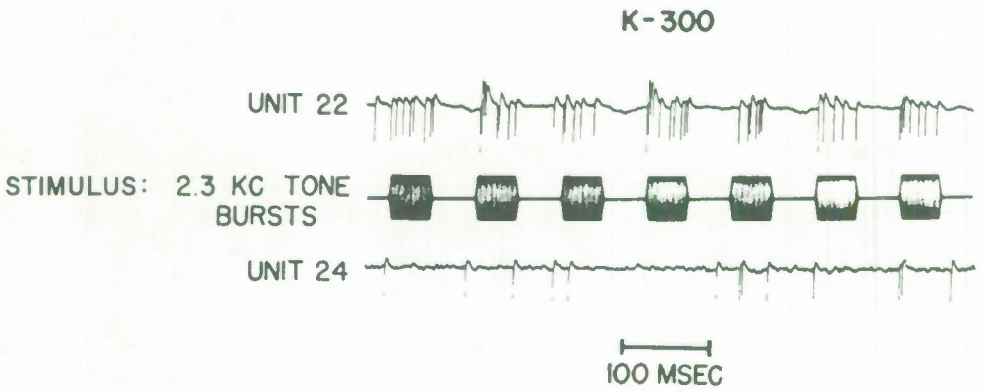

Fig. 4-1(b). Sample spike trains from two auditory nerve fibers of a cat. From Kiang (1965).

which occurs after the neuron has just discharged. Essentially, it represents a decreasing probability of neural discharge (per identical condition) with decreasing time since last discharge. For example, if the neuron is being stimulated by a repetitive stimulus, the ability of the neuron to discharge to each of the stimuli decreases as the repetition rate increases, all other things held constant. The minimal interpulse interval within which the neuron cannot discharge twice, regardless of stimulus intensity, is called the "least interval" or more commonly the "absolutely refractory period." The refractory period is then a short-term "memory effect" which is usually considered to be most effective in the few msec range. The effects of the refractory period will be evident in several of the histograms of interspike intervals shown below.

A further property of the temporal structure of neural outputs is a randomness which makes the times of neural discharge probabilistic rather than deterministic functions of the input to the neuron. Hence, in many cases, the discharges of neurons in response to a given signal can be described only with probability densities which are functions of the signal parameters.

Several other properties of the pulse trains from a neuron indicate that other memory-like processes are taking place. For example, adaptation is a term which often indicates a reduction in discharge rate during sustained stimulation; accommodation refers to a lack of excitation due to too gradual an increase in stimulus. The effects are most noticeable in the seconds range, but can often be detected for much longer times. We will discuss later similar effects that have been found in auditory nerve responses which are "memory-like" effects but occur even in the absence of discharge.

Properties of Auditory Nerve Fibers

The properties of the pulse trains generated by primary auditory neurons are closely related to the basilar membrane motion and the receptor action of the hair cells (or whatever else is acting), in addition to being subjected to the constraints of refractoriness and randomness (see Fig. 4-1b). These pulse trains are monitored by extremely fine electrodes placed within the auditory nerve. The dimensions of the electrode are such that electrical activity of single fibers (axons of individual neurons in the spiral ganglion) can be examined. Utilizing this technique, the electrophysiology of primary auditory neurons has been studied extensively although not exhaustively.

Spontaneous Activity

Each auditory nerve fiber conveys spike discharges to the brain even in the absence of any administered signal. These are called spontaneous discharges. The number of spike discharges per second varies from fiber to fiber from one every few seconds up to approximately one hundred per second. However, despite this wide range of average rates, the activities are remarkably similar if one looks at some properties of the distribution of intervals between spikes. For example, the mode of a histogram of interspike intervals is approximately 5-10 msec regardless of discharge rate (Fig. 4-2). In addition, the shapes of the histograms, which are a first-order approximation of the probability density of interspike interval times, are similar for all auditory nerve fibers. One such interval histogram is shown in Fig. 4-3. To a gross approximation, the decay after the mode is exponential. (Perhaps the decay farther away from the mode is truly exponential, but as the interval

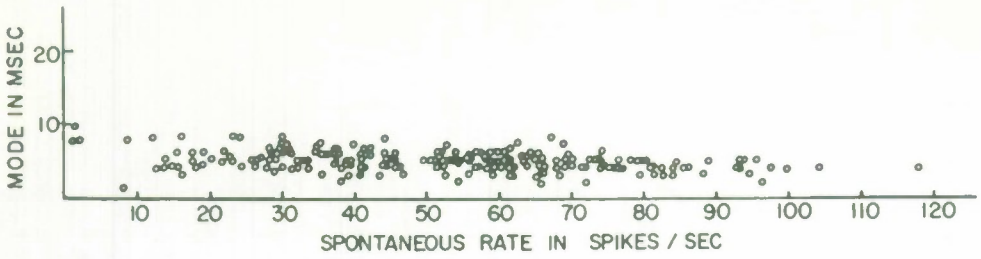

Fig. 4-2. The mode of the interval histograms of spontaneous activity plotted against rate of spontaneous discharges for 227 auditory nerve fibers. Each histogram is based on at least one minute of data. For units with low rates of spontaneous discharges, only long runs of data are meaningful. From Kiang (1965).

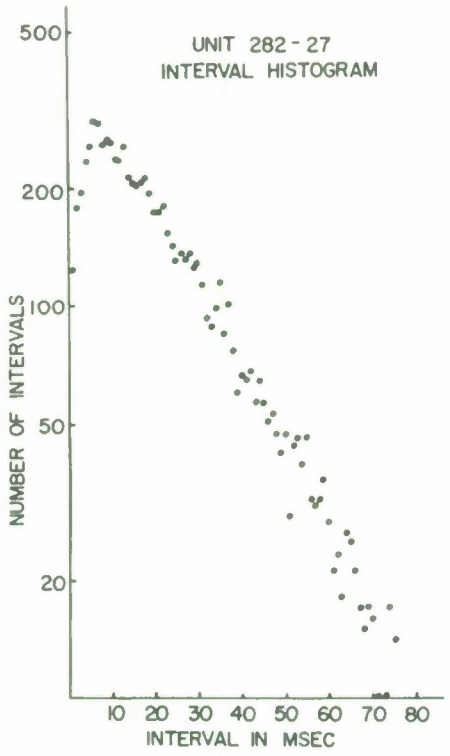

Fig. 4-3. Interval histogram of a unit with a medium rate of spontaneous discharges plotted on a semilogarithmic scale. The rate of spontaneous discharges was 45.5 spikes/sec. Three minutes of data are represented. CF: 0.58 kc. From Kiang (1965).

length increases, the size of the sample decreases and hence the statistical fluctuations do not allow an accurate assessment.) As a result, these spontaneous discharges have frequently been described as a Poisson process with a "dead time" that is a result of the refractory period. The perhaps remarkable feature is that the spontaneous activities of all of these fibers is qualitatively similar, thus presenting the picture that the primary neurons are a relatively homogeneous group of neurons.

Sinusoidal Stimulation

The responses of primary auditory neurons to several different types of acoustic stimulation also have been studied. A relative homogeneity among neurons also exists for their properties under stimulation. We have already seen that the motion of any region of the basilar membrane is frequency dependent (Fig. 3-18), and also that individual primary neurons are innervated by a limited region of the basilar membrane (Fig. 2-9). Hence, one can infer that only certain ranges of frequencies of stimulation should excite individual primary neurons. That this is the case is illustrated by the "tuning curves" of auditory nerve fibers shown in Fig. 4-4. In each case, the neurons respond by an increased rate of spike discharge, relative to the spontaneous rate, to signals whose frequency and intensity parameters lie within the boundaries of their tuning curves. The frequency of the most sensitive point of the tuning curve of a unit is the "characteristic frequency" (CF) or "best frequency" for that unit. As the signal level is increased, the neurons become responsive to a wider range of frequencies. This, of course, is related (but perhaps not directly) to the widening of the transfer function $H(s, x)$ for the basilar

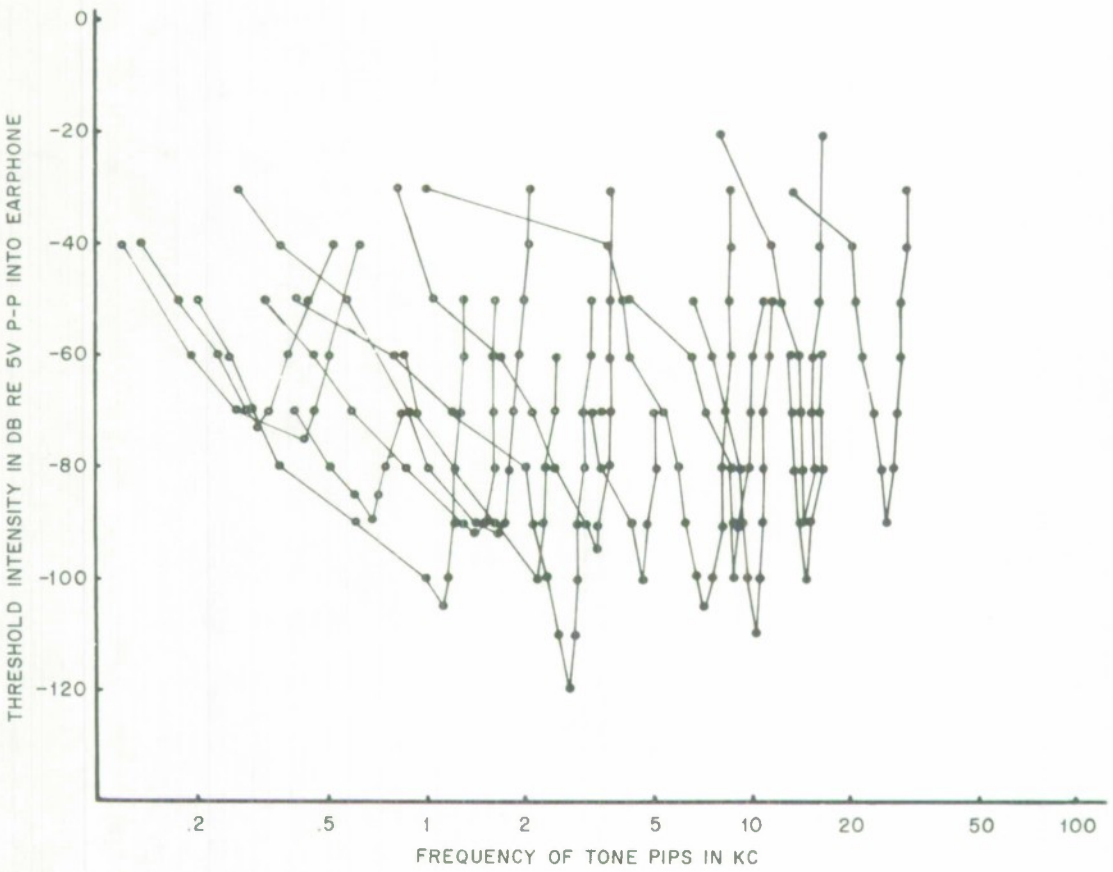

Fig. 4-4. Tuning curves for 16 different auditory nerve fibers in 9 cats. Each curve is obtained by setting the intensity of tone bursts and measuring the frequency range for which spike responses are obtained. The limits of this frequency range for a number of intensities are represented by points that have been joined to produce a "curve." The abscissa of the lowest point of each curve is defined as the characteristic frequency (CF or f_o) of that unit. The tone-burst stimuli had rise-fall times of 2.5 msec and a duration of 50 msec. From Kiang, et al. (1962).

membrane motion with increases in signal level (Fig. 3-21). The different CF's approximately relate to different positions of the primary neurons within the spiral ganglion as shown in Fig. 4-5.

For continuous sinusoidal stimulation by a signal whose parameters lie within the tuning curve, the distribution of spike discharges is similar to that of the spontaneous activity; that is, the interspike interval histograms have short modes and decay from the mode in an approximately exponential manner. The mode shifts slightly, but the major change is merely an increased rate of decay, resulting from a decrease in the number of long intervals due to the increase in the spike discharge rate. However, as one might expect, for cases of low-frequency stimulation (a few kHz or less) and over a wide range of signal levels, the spike discharges usually occur at preferred phase positions of the stimulus. In Fig. 4-6, the left-hand column shows histograms of spike discharge times relative to a fixed phase position of the stimulus. The right-hand column shows regular interspike-interval histograms of the same data. For lower frequencies, or higher resolution of computation or both, peaks appear in the interspike-interval histogram indicating that the spike discharges occur at intervals that are some multiple of the period of the sinusoidal stimulus.

Click Stimulations

One of the most widely used stimuli thus far in studies of the auditory nerve fibers have been acoustic clicks--usually generated by applying square electric pulses (0.1-msec long) to an earphone. The application of pulses of this dimension (termed clicks) is analogous to applying impulses to the system,

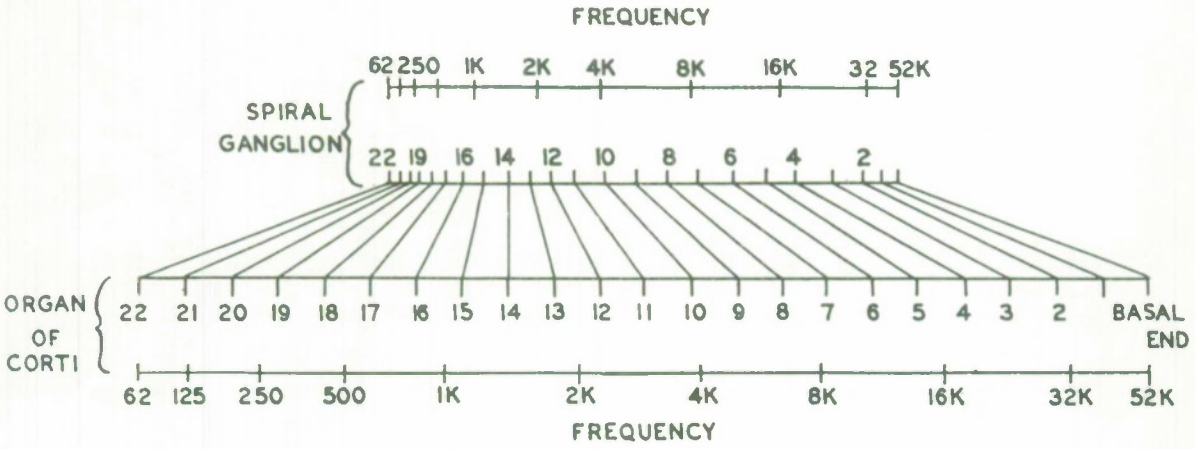

Fig. 4-5. Frequency has been spatially oriented along an organ of Corti of average length (22 mm). Anatomically related points and the spatial distribution of frequency on the spiral ganglion are also shown. It is interesting that, whereas the frequencies below 500 cps are located within the apical 20 percent of the organ of Corti, they are within the apical 10 percent of the spiral ganglion. From Schucknecht.

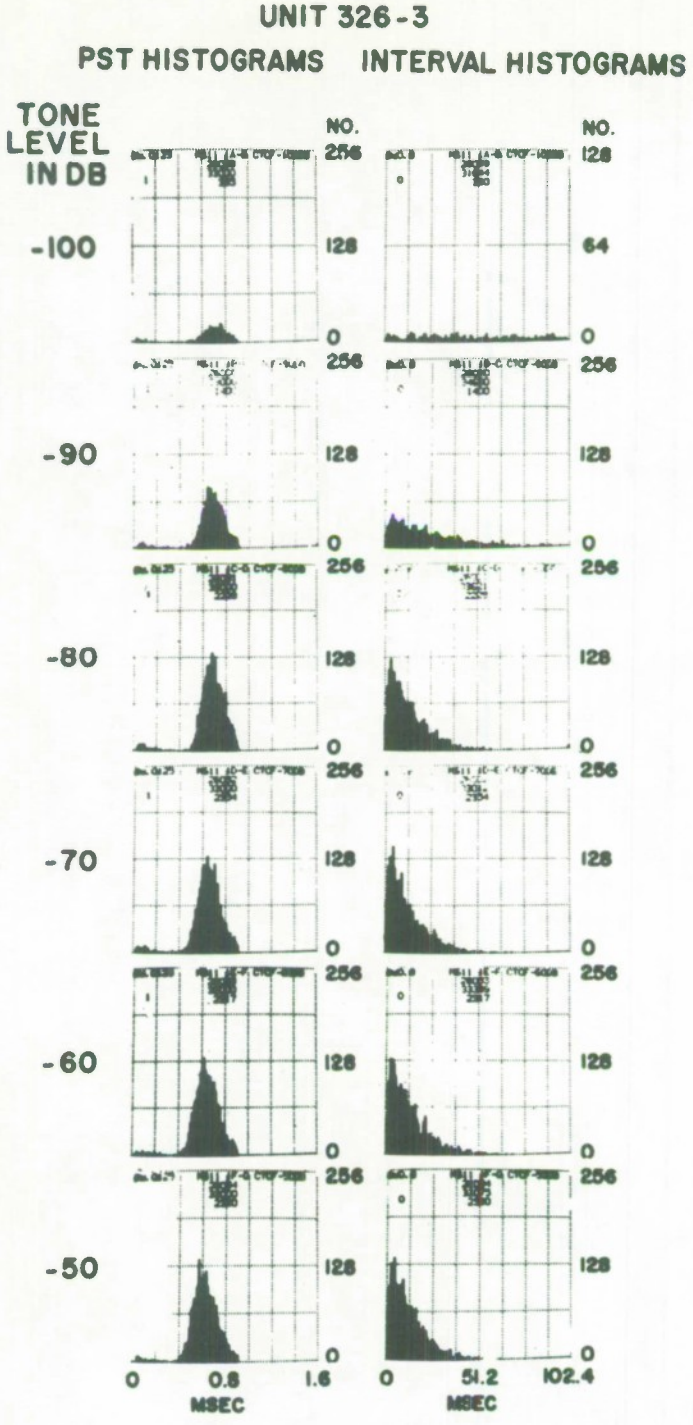

Fig. 4-6. PST and interval histograms of discharges from a single unit of the auditory nerve as the level of a continuous tone is changed. From Kiang (1965).

at least for frequencies of interest to us. The calculated impulse response described above of the basilar membrane (Fig. 3-28) then plays an important role in the expected response observed in the auditory nerve fibers. Figure 4-7 shows histograms of the number of spike discharges relative to the time of onset of stimulus (post-stimulus time [PST] histograms) for several auditory nerve fibers with different CF's (as measured with sinusoidal signals). The most obvious characteristic is the multiple peaks of the histograms for the units with CF's of approximately 4 kHz or less. It can be seen that the time interval between these peaks is approximately equal to 1/CF of each unit under question. That is, the spikes appear to be responding to the excursions of a rectified version of the impulse response of the position on the basilar membrane corresponding to that frequency. In addition, a decrease in the latency of response--time from signal presentation to first synchronized spike--is realized for units of increasing CF. This is in accord with the changes in propagation time of a basilar membrane disturbance as a function of distance from the oval window (Fig. 3-22). If the polarity of the clicks is reversed, thereby moving the stapes in an opposite direction, the peaks of the PST histogram are still 1/CF apart, but interleaved with those of the opposite polarity (Fig. 4-8).

At first glance, it appears from the above that there is a direct relationship between calculated basilar membrane displacement and auditory nerve fiber response. But upon finer examination, it can be seen that the number of peaks in the PST histograms of responses to clicks far exceeds the number of prominent peaks in the calculated basilar membrane impulse response (Fig. 3-28). A simple linear transformation from basilar membrane motion to

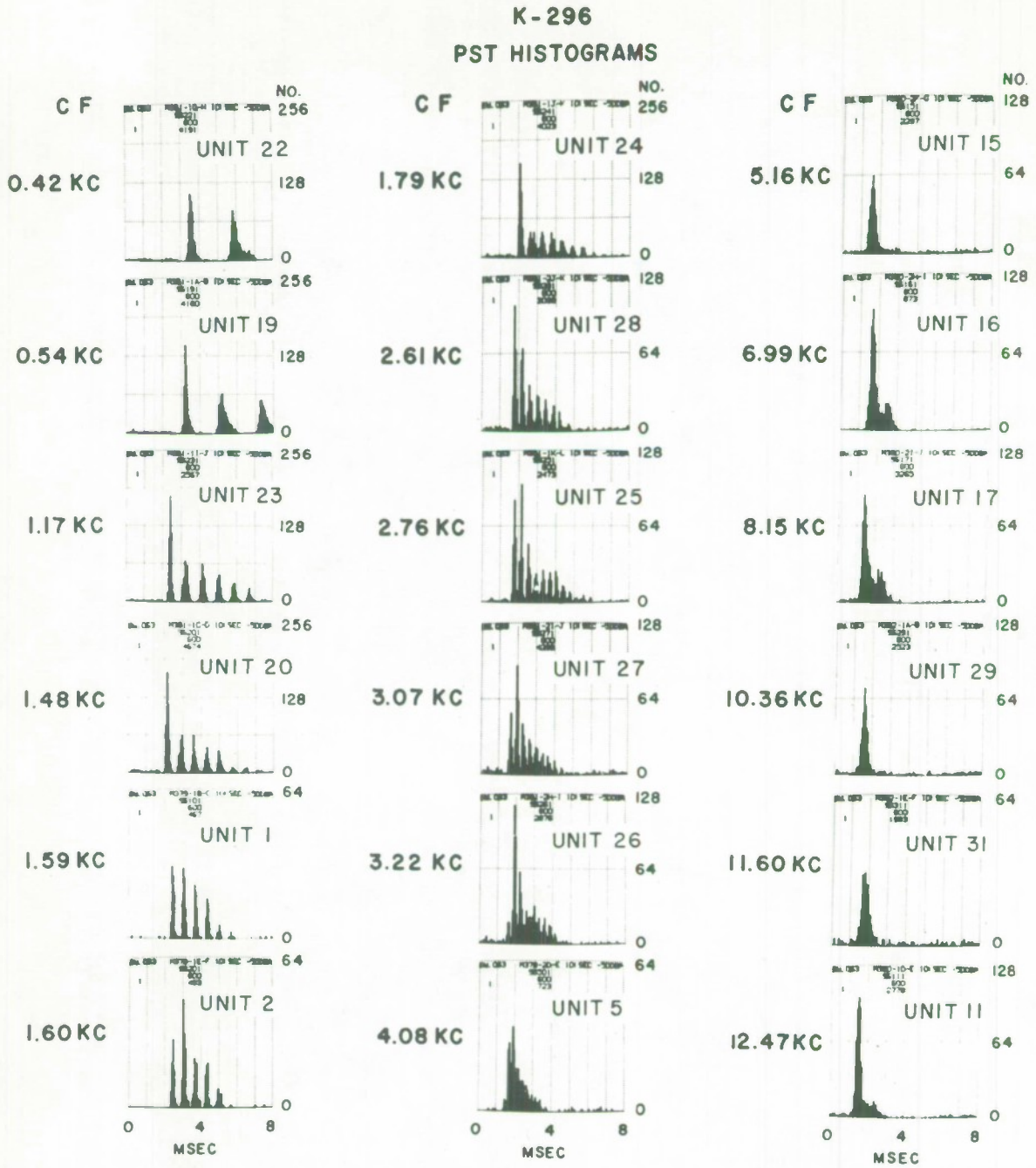

Fig. 4-7. PST histograms of responses to clicks from 18 units obtained in a single cat. The preparation was left undisturbed except when electrode position was changed. The histograms are arranged according to CF. Stimuli: 10/sec clicks. From Kiang (1965).

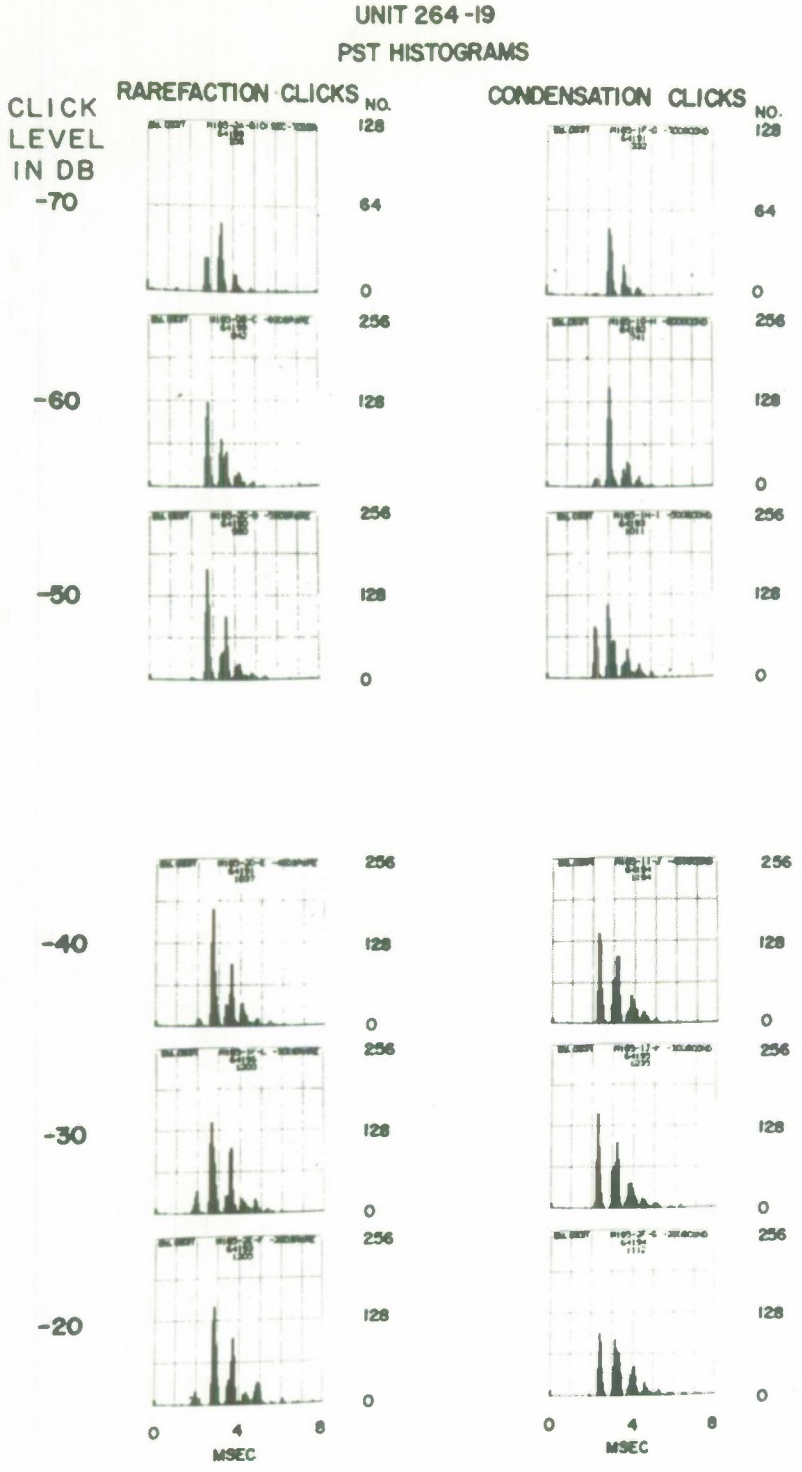

Fig. 4-8. Response patterns to rarefaction and condensation clicks as functions of click level. CF: 1.65 kc. Stimuli: 10/sec clicks. From Kiang (1965).

neural output does not appear to be adequate to describe the results of the auditory nerve fiber discharges.

Model of Auditory Nerve Discharges

We next examine an early modeling attempt by Weiss that was essentially based on the data thus far described for the primary neurons, the salient features of neurons in general, and the basilar membrane approximation of Flanagan.

A block diagram of Weiss's model and a diagram representing the neuron membrane potential variation is shown in Fig. 4-9. The model neuron--which was basically a threshold device with a built-in feedback loop to account for refractoriness--was driven by both a noise source, to simulate spontaneous activity and probabilistic behavior, and a signal proportional to the rectified motion of a single point on the basilar membrane. Weiss restricted his model to frequencies less than 2000 Hz so that the external and middle ears could be considered flat with frequency and, hence, ignored (Figs. 3-1, 3-4, 3-5).

With reference to Fig. 4-9, it can be seen that the noise input of Weiss's model is independent of the mechanical system and the transducer function G(y). The bandwidth of the noise input was an important factor, and was determined by trial and error until a reasonable match between interevent-interval histograms of the model to the actual interspike-interval histograms of auditory-nerve fibers was obtained. As a matter of interest, using a refractory-period time constant of approximately 1 msec, the cutoff frequencies of the noise passband were 2 and 5 kHz.

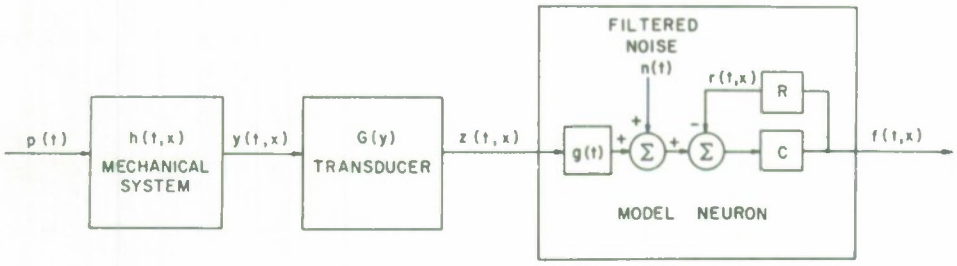

Fig. 4-9(a). Model relating the firing patterns of fibers in the auditory nerve to acoustic stimuli.

p(t)	pressure at the ear
y(t,x)	displacement of the cochlear partition
z(t,x)	output of a sensory cell
f(t,x)	sequence of spikes generated in an VIIIth-nerve fiber
r(t,x)	threshold potential
h(t,x)	impulse response of the mechanical system
G(y)	transducer function
g(t)	a linear filter
t	time
x	distance from the oval window to a point along the cochlear partition

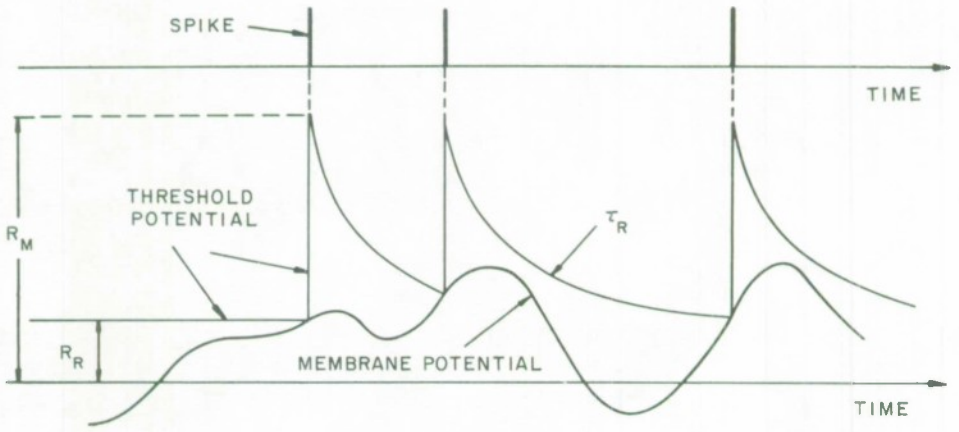

Fig. 4-9(b). Diagrammatic representation of membrane potential, threshold potential, and spike activity of the model neuron. From Weiss.

R_M maximum threshold potential

R_R resting threshold potential

τ_R time constant of the exponential decay of the threshold from its maximum to its resting value.

Coupling the mechanical motion of the basilar membrane to the model neuron necessitates a transducer (function). Weiss initially used a linear transducer function, $G(y) = y$. However, the nonconformity between responses of the model neurons under this condition to the responses of actual neurons was very evident. For example, when stimulated with clicks, the dynamic range of the model was not nearly that of a neuron, and also the proper number of peaks in a PST histogram of responses to clicks such as Fig. 4-7 failed to develop. Weiss reduced these discrepancies by introducing an ad hoc nonlinear transducer function, $G(y) = ay/(a + |y|)$, between basilar membrane motion and the input to his neuron model. With this latter scheme, Weiss was able to obtain better accord between the experimental and model data for the responses to click stimuli. In both cases, interleaving of peaks of PST histograms for clicks of opposite polarity was obtained.

The final version of Weiss's model generated data that approximated the experimental findings for spontaneous, sinusoidal, and click-stimulation situations. The agreements were qualitative in all cases, and reasonably quantitative for some. Some major disagreements did exist, e.g., the dynamic range of stimulation for the model with the nonlinear transducer function was too great when compared to the auditory nerve data, and the tuning curves of the model neurons were much wider than those measured from the nerve fibers.

A slight modification of Weiss's model was studied by Klatt. The frequency range was extended to include the entire range of the basilar membrane, and a different model was used for basilar membrane displacement (a resonance model of the cochlea similar to Fletcher's [see Chapter III]).

However, Klatt also was confronted with too few peaks in the responses to click stimulation, and with the problems of dynamic range and tuning curves that Weiss had. Instead of using a nonlinearity to modify the response, Klatt used a resonant circuit with a Q of approximately 10 between his neuron and point basilar membrane innervation. This introduced more peaks in the PST histogram and also narrowed the tuning curves.

The first attempts to model primary auditory neuron response were very tightly coupled to basilar membrane motion. In order to get reasonable model neuron responses, some ad hoc adjustments of the models were necessary (the nonlinearity by Weiss and the tuned circuit by Klatt). A complete assessment of the worth of these early models cannot be made yet because this work has not been continued.

Some Additional Properties of Auditory Nerve Fibers

An obvious over-simplification in the above models was that of a single point connection between the model neuron and basilar membrane. The anatomical connections of the fibers that exist in the organ of Corti strongly indicate that a zone or region of the basilar membrane contributes to the effective stimulus that excites a primary neuron.

An electrophysiological property that reflects this zone of stimulation is that of inhibition. The rate of auditory nerve discharges in response to sinusoidal stimuli whose parameters are within the tuning curve can be lessened, and in cases reduced to the rate of the spontaneous discharges, by applying a second sinusoidal signal. That is, a second tone can cancel or inhibit the response of a fiber to a first tone. The "inhibitory areas," i. e., regions on

the frequency, signal-level plane in which a second sinusoidal signal must lie to inhibit the responses to a first tone are functions of the parameters of the first tone. In general, these regions lie adjacent to the tuning curve (although they can overlap), both on the high- and low-frequency sides, and can be quite extensive in area.

Most of these inhibitory properties, as monitored in auditory nerve fibers, have only recently been quantitatively studied in detail. (Thus, certain properties of inhibition are still somewhat controversial, e.g., whether or not a single tone can "inhibit" spontaneous discharges.) As a result, most models of the peripheral auditory system to date have not incorporated a mechanism for inhibitory action. It appears that a point connection to the basilar membrane is not adequate to account for inhibition, except perhaps in some special cases at low frequencies where certain opportune phasing may take place. The recent work of Sachs on quantitative aspects of inhibition with two-tone stimuli may suggest a realistic model for the interconnection between primary fibers and basilar membrane.

Furman has recently investigated the characteristics of a variety of possible inhibitory connection schemes among elements lying within the cochlea. He was motivated by the observations from experimental data that the tuning curves of primary neurons are considerably sharper than would be expected if these functions were linearly dependent on displacement of the basilar membrane at a single point. He intended to demonstrate that the tuning curves of neurons could be sharpened by means of inhibitory interaction. At the same time, he attempted to obtain a nonlinear relation between neuron sensitivity and basilar membrane motion.

Although one modeling scheme appeared to "sharpen" response areas, it did not generate a nonlinearity, while another generated a nonlinearity but did not sharpen the response area. Nevertheless, Furman used the latter scheme in a simulation of the bat peripheral auditory system as shown in Fig. 4-10. This is the most complete model described yet since it includes the peripheral, middle, and inner ears, including a scheme for inter-unit coupling, and a neuron model. Furman demonstrated for this scheme a qualitative agreement for (1) temporal configuration of spikes elicited by stimulus bursts; (2) spontaneous spike configurations; and (3) two-tone inhibition as detected in bats.

Further attempts to find other means of determining what is occurring between membrane motion and neural excitation are being applied to the data on the auditory nerve fiber discharges. Gray has carried out a statistical analysis of auditory nerve fiber data utilizing processing techniques which eliminate the obscuring effects of refractoriness* in some forms of processed data such as PST histograms. Since the refractoriness is due to inherent properties of the neuron, the resultant data more closely reflect the "effective" stimulus to the neuron. His results have shown that even the "recovered" discharges are not linearly related to basilar membrane motion or any linear transformation of it. In fact, his results indicate that,in addition to considering refractoriness in a model neuron, one must include an additional state variable

* He essentially examines only those discharges which occur after the neuron is in a "recovered" state. Roughly speaking, the recovered state has been demonstrated to be any time after about 25 msec from the just previous discharge, and it represents a state of the neuron when the probability of firing is no longer a function of the time since last firing.

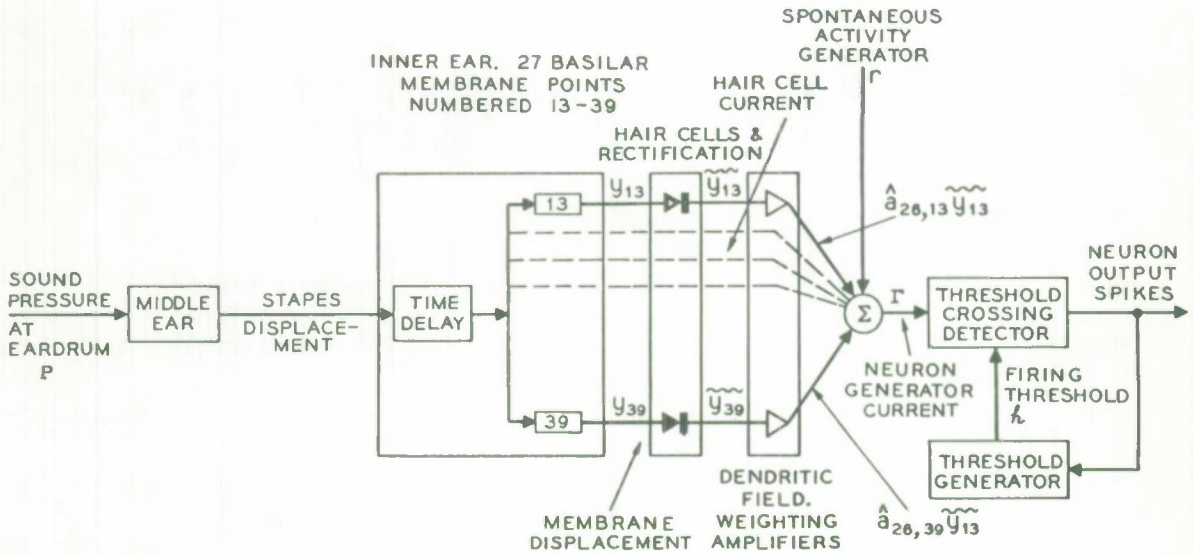

Fig. 4-10. Model of peripheral auditory system of a bat. From Furman.

which is affected by the stimulus even in the absence of neuron discharge. He calls this a "depletion effect" because recovered probabilities of discharge diminish after moderate to large values of recovered probability have occurred--even when the neuron does not discharge. Gray states in summary "...the occurrence of large recovered probabilities would affect this variable in such a way that recovered probabilities resulting from subsequent mechanical motion would be diminished." To date no neural model for the peripheral auditory system has incorporated this additional state variable. But it presumably would play an important part in the character of an overall model of the periphery. These findings, of course, are very recent and quantitative details of a large population of neurons are not yet available. Thus the mathematical framework for such a model is not yet developed.

Summary

1. A primary neuron can be considered to be a pulse generator whose output temporal structure is governed by basilar membrane motion, intrinsic properties of neurons, and a noise component.

2. Large quantities of experimental data are available and indicate that all of the primary fibers are qualitatively similar in their properties, i.e., the temporal structure of their spike train.

3. Several attempts to model these data have yielded qualitative agreement, but all have failed in some quantitative sense mainly due to the lack of facts pertinent to the functions occurring between basilar membrane motion and neuron output.

4. Recent, but not yet complete, studies pertaining to inhibition properties and to results of new processing techniques applied to experimental data appear promising in attempts to obtain a more realistic model of the peripheral system up to the outputs of the primary neurons.

Organ of Corti, Transduction, and Inhibition References

L. A. Bicking, "Some Quantitative Studies on Retinal Ganglion Cells," Ph. D. Dissertation, Johns Hopkins University (1965).

W. van Bergeijk, "Studies with Artificial Neurons, II: Analog of the External Spiral Innervation of the Cochlea," Kybernetics 3, 102-107 (1962).

H. Davis, "Transmission and Transduction in the Cochlea," The Laryngoscope 68, 359-382 (1958).

G. G. Furman, "Cochlear Sensory-Field Effects Based on Inter-Unit Coupling," Report No. 65-19, Electronics Research Laboratory, University of California, Berkeley, California.

G. G. Furman and L. S. Frishkopf, "Model of Neural Inhibition in the Mammalian Cochlea," J. Acoust. Soc. Am. 36, 2194-2201 (1964).

H. K. Hartline, F. Ratcliff, and W. H. Miller, "Inhibitory Interaction in the Retina and its Significance in Vision," Nervous Inhibition (Pergammon Press, N. Y., 1961).

I. Sando, "The Anatomical Interrelationships of the Cochlear Nerve Fibers," Acta Oro-laryng. 59 (1965).

H. F. Schuknecht, "Neuroanatomical Correlates of Auditory Sensitivity and Pitch Discrimination in the Cat," Neural Mechanisms of the Auditory and Vestibular Systems, ed. by G. L. Rasmussen and W. F. Windle, pp. 76-90 (Charles C. Thomas · Publisher, 1960).

M. B. Sachs, "Auditory Nerve Fiber Responses to Two-Tone Stimuli," Ph. D. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts (August 1966).

Auditory Nerve References

G. L. Gerstein and N. Y-S. Kiang, "An Approach to the Quantitative Analysis of Electrophysiological Data from Single Neurons," Biosphy. J. 1, 15-28 (1960).

P. R. Gray, "A Statistical Analysis of Electrophysiological Data from Auditory Nerve Fibers in Cat," Ph. D. Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts (1965).

Auditory Nerve References (Concl.)

N. Y-S. Kiang, T. Watanabe, E. C. Thomas, and L. F. Clark, "Stimulus Coding in the Cat's Auditory Nerve," Ann. Otol. Rhinol-Laryngol. 71, 1009-1027 (1962).

N. Y-S. Kiang, with the assistance of T. Watanabe, E. C. Thomas and L. F. Clark, Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, M. I. T. Monograph No. 35 (M. I. T. Press, Cambridge, Massachusetts, 1965).

A. Rupert, G. Moushegian, and R. Galambos, "Unit Responses to Sound from Auditory Nerve of the Cat," J. Neurophysical 26, 449-465 (1963).

W. M. Seibert, "Some Implications of the Stochastic Behavior of Primary Auditory Neurons," Kybernetik 2, 206-215 (1965).

T. F. Weiss, "A Model for Firing Patterns of Auditory Nerve Fibers," R. L. E., Massachusetts Institute of Technology Technical Report No. 418 (1964).

D. H. Klatt, "Theories of Aural Physiology," Report No. 13, Communication Sciences Laboratory, University of Michigan, Ann Arbor, Michigan (November 1964).

J. L. Flanagan, Speech Analysis, Synthesis, and Perception (Academic Press, New York, 1965).

M. A. B. Brazier, The Electrical Activity of the Nervous System (The MacMillan Company, New York, 1958).

CHAPTER V

Summary and Conclusions of the Study

A. REVIEW OF THE STATUS OF PERIPHERAL AUDITORY SUBSYSTEMS

External Ear

The pertinent data with respect to sound transmission in the external ear canal of humans are available. These data indicate that the external ear could be satisfactorily modeled as a lossless transmission line terminated in a load impedance corresponding to the input impedance of the middle ear as seen at the eardrum. There are no efferent or feedback mechanisms to consider in characterizing the external ear.

Middle Ear

The most complete data on the middle ear transfer function were taken on cats rather than humans. These data allow a separation of the effects of the ossicles and of the middle ear cavities. Sound transmission via the ossicles alone has been satisfactorily modeled by a lumped element circuit.

Although corresponding data do not exist for the human middle ear, a satisfactory model for the ossicles alone would again be a lumped element circuit of the same form as that developed for the cat. When the pertinent data become available, they will serve to determine the values of the circuit elements specifically for humans. Temporarily, we can use as the circuit model for the human ossicles that model developed for cat ossicles, since we expect these circuits to be quite similar in detail.

On the other hand, there is a considerable difference between the middle ear cavities of the cat and those of humans. We expect the effects of the human middle ear cavities on sound transmission through the middle ear to be less significant than in the cat. Until better data exist, we choose to neglect the effects of the middle ear cavities in modeling the human middle ear transfer function.

The question of efferents cannot be entirely neglected in the treatment of the middle ear. However, it is believed that efferent control of the middle ear transfer function (by means of the tensor tympani and stapedius muscles) is reflexive in nature and serves mainly as a sort of automatic gain control to protect the cochlea from loud sounds. If this were the case, efferents would not play an important role in the intensity range of normal speech sounds, but the data are meager on this subject.

Cochlea

The pertinent experimental data on cochlear mechanics is due almost entirely to the work of Georg von Bekesy. Cochlear models based on his data are of two fundamentally different types--resonance models and dispersion models. A clear discrimination as to which type of model more closely resembles the actual cochlear action cannot be made at this time because of the incomplete knowledge of some cochlea parameters as well as some aspects of sound transmission within the cochlea. Since the data are incomplete, about all that one can do in modeling the cochlea is to match as many model characteristics to experimental data as possible. Both resonance and dispersion models of the cochlea have been presented and appear to be adequate to model many of the details of cochlear mechanics.

Our own work on cochlear modeling is incomplete although it has progressed beyond that reported here. We hope to develop both resonance and dispersion models for detailed comparison of characteristics. This work will be reported at a later time.

The fine motions that the elements of the organ of Corti undergo during sound stimulation are qualitatively known. Since it appears that these fine motions are memoryless (i. e., they depend only on the instantaneous basilar membrane deflection pattern), they should not be too difficult to include in a model of the cochlea if such detail is needed.

There appear to be no mechanical efferents (via muscles) that affect sound transmission in the cochlea.

Mechanical-to-Neural Transduction

Although it appears quite feasible to include the fine motions of the organ of Corti in models of the cochlea, the main problem in doing so is that the precise features of these fine motions that govern stimulation of the hair cells are really not yet known. To compound the problem, it is reasonable to expect that the inner and outer rows of hair cells are stimulated differently, but this has never been clearly demonstrated. This dilemma arises out of a gap in experimental data. The electrophysiology of a hair cell has not been experimentally studied under acoustic stimulation of the cochlea. Single hair-cell outputs are not available either. As discussed earlier, the primary fiber data are all that are available from which to infer details of the transduction process. But since the plan of interconnections of hair cells to primary neurons is also largely unknown, it is an almost impossible task to understand

transduction at present. The so-called two-tone (or multiple tone) inhibition effect in primary fibers is currently under investigation and should yield very basic information concerning the plan of innervation of primaries.

The efferent bundle of the auditory nerve consists of only about 500 fibers which terminate on the hair cells themselves. But there are many more than this number of inhibitory synapses on the hair cells, indicating that each efferent fiber innervates many hair cells. The effect of these efferents on hair cell function is not known. The effect of electrical stimulation of the efferent bundle on primary neuron outputs is currently being investigated. Normal functioning of the efferent bundle may not be observable in anesthetized cats because of the effect of anesthesia on the auditory centers in the brain where the efferents originate.

In summary, the details of the mechanical-to-neural transduction process are not understood at present. The effect of efferents on transduction and the plan of innervation of primary fibers are the objects of current experimental investigations. Thus, it is impossible to predict when in the future the transduction process will be well enough understood to allow realistic modeling and simulation.

Primary Auditory Neurons

A great amount of experimental data has been taken on single primary neuron outputs but only for a small class of stimuli (clicks, single tones and tone bursts, and, more recently, multiple tones). These data indicate a homogeneity among primary neurons, which should ease the modeling problem. However, since the details of the transduction process are not

yet known, the inputs to primary neurons are not really known, making neuron modeling difficult in spite of the abundance of primary fiber data.

A neuron model proposed by Weiss incorporated many of the basic, characteristics of neurons, such as spontaneous activity, refractoriness, and a nonlinear relation between stimulus intensity and firing rate. This model, when driven by the displacement of a single point on the basilar membrane, gave good qualitative agreement with much of the primary fiber data. The more recent investigations of Gray have indicated that more complexity is required in modeling neurons. One disturbing finding by Gray was that primary neurons with similar firing histories did not exhibit a linear relation between stimulus intensity and probability of firing in an interval. Thus at this time there is no general agreement on a primary neuron model. There are no known efferent connections directly to the primary neurons.

The Central Auditory System

The cochlear nucleus contains all of the auditory nerve fiber terminations but, in addition, is innervated by many higher auditory centers via efferent pathways. It is more complicated than the hair cell--primary neuron--efferent system discussed above. The general function of the cochlear nuclei are not presently known. Normal functioning of the cochlear nucleus neurons may not be observable in anesthetized animals because of the effects of anesthesia on many of the inputs from the higher auditory centers. The function of parts beyond the peripheral auditory system is thus clearly beyond our understanding at present.

B. STATUS OF PERIPHERAL AUDITORY SYSTEM MODELS

Since the mechanical or sound-conducting portions of the auditory system are linear and are believed to have no important efferent connections as far as speech sounds are concerned, there are no complications introduced by combining the models of these subsystems to obtain a transfer function characterizing the entire mechanical system. It is, therefore, possible to accurately relate the sound stimulus at the entrance to the ear canal to the actual inputs to the transducer. However, modeling the overall system can proceed no further since the details of the next function, that of mechanical-to-neural transduction, are not known. Moreover, efferents or feedback paths to the transducer may be of critical importance and they also are not well enough known to model. Nevertheless, the temptation to extend the system model to the neural levels has been great, and this has led to some speculation about the transduction process for modeling purposes, with interesting results.

Perhaps the simplest and most tempting speculation to make is that the excitation of a single primary fiber depends only on the instantaneous displacement of a single point on the basilar membrane. This is tantamount to assuming that all the hair cells that innervate a particular primary neuron are closely spaced along the cochlea. From the PST histograms for acoustic clicks, it seems very likely at first glance that the probability of a primary neuron firing in some interval could be related to the rectified displacement of the basilar membrane (only displacements in one direction from equilibrium seem to cause firings). Changing the polarity of the acoustic clicks from rarefaction to compression produces PST histogram peaks which interleave

with each other, as would be the case with the peaks of the rectified basilar membrane displacement impulse response. However, closer study of this model reveals some serious shortcomings.

1. Primary fiber PST histograms for acoustic clicks have peaks that are spaced 1/CF seconds apart, where CF is the fiber's characteristic frequency. Fibers have been found with PST histogram peak spacing corresponding to frequencies as high as 4 kHz. The combined impulse response of the middle ear and basilar membrane, however, will resemble the middle ear impulse response at such high frequency points on the basilar membrane due to the low pass, highly damped response of the middle ear. The measured middle ear and basilar membrane frequency responses, in which we have considerable confidence, are inconsistent with this conjecture of single-point excitation of single fibers.

2. The curves of constant firing rate in the stimulus frequency-stimulus intensity plane for a primary neuron with continuous tone excitation have significantly narrower bandwidth than the basilar membrane frequency responses corresponding to the same best frequency as the neuron. At least for high frequency primaries, one would expect that such iso-rate curves correspond to a fixed level of excitation. For instance, in Weiss's neuron model, the shape of any iso-rate curve would be independent of the nonlinearity G(y). Iso-rate curves should, therefore, have the same shape as the basilar membrane frequency responses in the case of single-point excitation.

3. It has been demonstrated that the response of a primary neuron to a continuous tone can be suppressed in varying degrees by a second

tone (the so-called two-tone primary inhibition). No linear functional of displacement of a point on the basilar membrane can yield an excitation function consistent with this inhibition because of the principle of superposition for linear systems.

C. THE ROLE OF COMPUTER SIMULATION

Computer simulation of models has already proved to be a valuable tool in modeling the peripheral auditory system. For instance,

1. simulation of the mechanical subsystems was necessary to obtain the driving function for neural models;

2. comparison of neural model outputs with experimental data has pointed out deficiencies in the neural model;

3. in the future, simulation should be valuable in designing new experiments to further test promising models and also to allow correlation of model performance and psychoacoustic experiments.

There is one further point regarding computer simulation of auditory system models, or any neural system models for that matter. The gap in experimental data between mechanical motions of the organ of Corti and primary neuron outputs makes it difficult to arrive at models for the neural stages of the peripheral auditory system. The difficulties in obtaining data to fill this gap may never be totally surmounted. A similar dilemma is also obvious when one tries to understand the function of the cochlea nucleus and higher auditory centers of the brain by studying the available experimental data. In such situations where complete experimental data may not be

obtainable, it seems that perhaps the only way of arriving at models for neural processing may be sagacious guesses of the form of models and comparison of model performance obtained by computer simulation with whatever data are available.

Acknowledgment

The authors consulted with many people working in a diversity of specialties during the course of this study. Several deserve special acknowledgment. They are:

Professor William T. Peake of M. I. T. for helpful discussions of the middle ear function as well as the use of his middle ear circuit model.

Mrs. Maxine H. Simon for writing the computer programs and providing the computations for our modeling efforts.

Dr. Barney Reiffen for valuable criticisms during the course of many discussions.

DOCUMENT CONTROL DATA – R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Lincoln Laboratory, M.I.T.	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None

3. REPORT TITLE

Signal Processing Characteristics of the Peripheral Auditory System

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Note

5. AUTHOR(S) *(Last name, first name, initial)*

Goblick, Thomas J., Jr. and Pfeiffer, Russell R.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
30 September 1966	146	70

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF 19(628)-5167 b. PROJECT NO. 649L c. d.	Technical Note 1966-50 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* ESD-TR-66-449

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Air Force Systems Command, USAF

13. ABSTRACT

The fundamental question in speech compression is that of determining the minimum information rate that must be maintained between speaker and listener in order to achieve a specified level of speech fidelity or quality. The problem in answering this question is that a measure of speech quality must first be defined. Any meaningful definition of speech quality clearly must consider the manner in which speech is processed by the listener. If the details of the signal processing in the auditory system were known, speech quality could be defined in terms of the sensitivity of the listener to distortions of signals within the auditory system.

A study of the manner in which sounds are processed by the human auditory system was done to provide the basic information to define a measure of speech fidelity. The mechanical or sound-conducting parts of the auditory system are reasonably well understood and can be considered as linear systems in an engineering sense. The neural processing of the peripheral auditory system is only partly understood. Further experimental work is necessary.

14. KEY WORDS

signal processing characteristics
auditory signals
peripheral auditory system
speech
speech compression
speech recognition